Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

# JOURNAL OF GENETIC PSYCHOLOGY

## Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

JOHN E. ANDERSON
University of Minnesota

S. K. BOSE
University College of Science, Calcutta, India

CHARLOTTE BÜHLER
University of Southern California

CYRIL BURT
University of London

LEONARD CARMICHAEL
Smithsonian Institution

SANTE DE SANCTIS
R. Università di Roma

ARNOLD GESELL
Yale University

WILLIAM HEALY
Judge Baker Foundation, Boston

ARTHUR T. JERSILD
Teachers College, Columbia University

BUFORD JOHNSON
The Johns Hopkins University

HAROLD E. JONES
University of California

TRUMAN L. KELLEY
Harvard University

K. S. LASHLEY
Harvard University

H. P. MAITI
B. M. Institute of Child Development Ahmedabad, India

YOSHIKAZU OHWAKI
Tohoku University, Japan

HENRI PIÉRON
Université de Paris

ROBERT R. SEARS
Stanford University

GEORGE D. STODDARD
University of Illinois

LEWIS M. TERMAN
Stanford University

C. J. WARDEN
Columbia University

JOHN B. WATSON
New York City

TATSURO YATABE
Kyoto University, Japan

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

## VOLUME 90

1957

Copyright, 1957, by The Journal Press

Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

# INDEX OF AUTHORS

Adcock, Cyril J. .......... 121
Allen, Robert M. .......... 280
Bayley, Nancy .......... 3
Boguslavsky, George W. .......... 83
Brown, W. Lynn .......... 33, 89, 97, 259
Brownstein, Aaron J. .......... 213
Carey, Janice E. .......... 69
Fenn, James D. .......... 59
Fosmire, Frederick R. .......... 33
Ghiselli, Edwin E. .......... 131
Goss, Albert E. .......... 59, 69
Green, Robert C., Jr. .......... 267
Halas, Edward S. .......... 89, 97, 259
Iwahara, Shinkuro .......... 213
Kenny, Douglas T. .......... 75
King, John A. .......... 151
Klingberg, Göte .......... 227
Levinson, Boris M. .......... 17
Mann, Helene Powner .......... 167
Marx, Melvin H. .......... 213, 219
Messick, Samuel J. .......... 23
Moyer, K. E. .......... 103
Salzinger, Kurt .......... 239
Seeman, William .......... 267
Smith, Madorah E. .......... 143, 255
Solley, Charles M. .......... 23
Staudt, Virginia M. .......... 271
Steinschneider, Alfred .......... 109
Stern, John A. .......... 203
Wisner, Robert J. .......... 75

Bureau Edn. Research
Dated 3.9.70
Acc. No. J-312

# TABLE OF CONTENTS

Data on the growth of intelligence between 16 and 21 years as measured by the Wechsler-Bellevue Scale . . . . . . . . . . 3
By Nancy Bayley

A comparative study of the intelligence of Jewish preschool boys and girls of orthodox parentage . . . . . . . . . . 17
By Boris M. Levinson

Probability learning in children: Some exploratory studies . . . . . 23
By Samuel J. Messick and Charles M. Solley

A study of spatial learning in a linear maze . . . . . . . . 33
By Frederick R. Fosmire and W. Lynn Brown

The rôle of mediating verbal responses in the conceptual sorting behavior of normals and paranoid schizophrenics . . . . . . . 59
By James D. Fenn and Albert E. Goss

The rôle of mediating verbal responses in the conceptual sorting behavior of children . . . . . . . . . . . . . . . 69
By Janice E. Carey and Albert E. Goss

An evaluation of Shock's "growth curves" . . . . . . . . . 75
By Robert J. Wisner and Douglas T. Kenny

Adaptive responses and the contiguity principle . . . . . . . 83
By George W. Boguslavsky

Terminal extinction in a multiple-$T$ maze within a heterogeneous environment 89
By W. Lynn Brown and Edward S. Halas

Terminal extinction in a multiple-$T$ maze within a homogeneous environment 97
By W. Lynn Brown and Edward S. Halas

Relationship between emotional elimination and the persistence of an anxiety motivated response . . . . . . . . . . . . 103
By K. E. Moyer

Audiogenic seizures: Insulin and high blood sugar . . . . . . . 109
By Alfred Steinschneider

BOOKS . . . . . . . . . . . . . . . . . 119

CRITICAL REVIEWS OF RECENT BOOKS . . . . . . . . 121

Allport, Gordon W. *Becoming.* (Reviewed by Cyril J. Adcock) . . 121

BOOKS RECENTLY RECEIVED . . . . . . . . . . 127

The relationship between intelligence and age among superior adults . . 131
By Edwin E. Ghiselli

Word variety as a measure of bilingualism in preschool children . . . 143
By Madorah E. Smith

Relationships between early social experience and adult aggressive behavior in inbred mice . . . . . . . . . . . . . 151
By John A. King

Some hypotheses on perceptual and learning processes with their applications to the process of reading: A preliminary note . . . . . . 167
BY HELENE POWNER MANN

The effect of frontal cortical lesions on activity wheel and open-field behavior 203
BY JOHN A. STERN

Hoarding behavior in the hooded rat as a function of varied alley illumination . . . . . . . . . . . . . . 213
BY MELVIN H. MARX, SHINKURO IWAHARA, AND AARON J. BROWNSTEIN

Spread of effect: Animal studies . . . . . . . . . . 219
BY MELVIN H. MARX

The distinction between living and not living among 7-10-year-old children, with some remarks concerning the so-called animism controversy . . 227
BY GOTE KLINGBERG

Academic achievement in a group of mentally disturbed adolescents in a residential treatment setting . . . . . . . . . 239
BY KURT SALZINGER

Progress in the use of English after 22 years by children of Chinese ancestry in Honolulu . . . . . . . . . . . 255
BY MADORAH E. SMITH

Latent extinction in a multiple-*T* Maze within heterogeneous and homogeneous environments . . . . . . . . . . . 259
BY W. LYNN BROWN AND EDWARD S. HALAS

SHORT ARTICLES AND NOTES . . . . . . . . . . 267

Drive in illness: An experimental study of barrier behavior of (alloxan) diabetic and nondiabetic animals (William Seeman and Robert C. Green, Jr.) . . . . . . . . . . . 267

Graduate schools ought to train teachers of psychology (Virginia M. Staudt) . . . . . . . . . . . 271

BOOKS . . . . . . . . . . . . . . 275

BOOKS RECENTLY PUBLISHED . . . . . . . . . 275

Henry, W. E. *The Analysis of Fantasy: The Thematic Apperception Technique in the Study of Personality*. (Reviewed by Robert M. Allen) . . . . . . . . . . . 280

*The Journal of Genetic Psychology*, 1957, **90**, 3-15.

# DATA ON THE GROWTH OF INTELLIGENCE BETWEEN 16 AND 21 YEARS AS MEASURED BY THE WECHSLER-BELLEVUE SCALE*

*Institute of Child Welfare, University of California*

---

NANCY BAYLEY[1]

---

## A. INTRODUCTION

How long do individuals continue to grow in intelligence? Do some persons reach their adult intellectual capacities sooner than others? And if so, what are the differentiating characteristics of those whose intelligence continues to grow for a longer, as compared with a shorter time? Do some intellectual functions mature earlier than others? These are some of the questions raised by those who are concerned with the theory of intellectual development.

There is general agreement on the average trends of mental growth, among such studies as those of Jones and Conrad (7), of Miles (8), and of Wechsler (12), in this country, and of Raven (10) and Foulds (3) in England. In all of these studies the mean scores for age, on tests of intelligence, increase to about 21 years of age, after which there is a gradual decline in scores earned in most types of ability. There is, furthermore, ample evidence that different types of mental functions increase and decline at somewhat differing rates. These studies, however, have been made on cross-sectional samples, and tell us only of the average trends for age of people tested at the same time.

Ideally, in order to find the answers to our questions, we should have repeated tests of a wide variety of intellectual functions, given at intervals of a year or two, on relatively unselected samples, and continued over periods extending from childhood well into adult years. This ideal is difficult to meet, so we must piece together such evidence as we can get, in any series where tests have been repeated for some of the relevant ages.

A few longitudinal studies, including ages extending into the early twenties, have been made in which the subjects were tested at least twice. Free-

---

*Received in the Editorial Office on March 4, 1954, but delayed in publication because of the fire.

[1] Since October, 1954, Laboratory of Psychology, National Institute of Mental Health, Bethesda, Maryland.

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated .......
Accs No .......

man and Flory (4), for example, were able to get later tests in college of 26 of the subjects in the University of Chicago longitudinal study. R. L. Thorndike (11) made a statistical study of 1,004 college students who had taken American Council on Education Psychological Examination at least twice, at annual intervals, for ages ranging from 13½ to 20 years. In both of these studies the subjects' scores continued to increase to the oldest age tested, which ranged up to 20 and 21 years.

Both of these samples, however, are exclusively college students; their continued education might be a factor in causing their scores to increase. What is more, the likely influence of education on scores is facilitated by the fact that in both studies the tests used are group tests, with a strong weighting of verbal and mathematical concepts. On the other hand, Freeman and Flory's data for ages up to 16 give some indication that children with low scores may at that age be growing more rapidly than the high scorers.

## B. Sample and Program of Tests

The present report is offered as one small addition to the available test data. Although it still deals with a selected sample, it is not confined exclusively to the highly intelligent nor to the college educated. Furthermore the test under consideration here is an individual test which yields both verbal and performance scores.

The children of the Berkeley Growth Study (1, 6) took the Wechsler-Bellevue Adult Intelligence Scale (Form I) at 16, 18, and 21 years. All of the tests were administered by the author. Thirty-three of the subjects were tested at all three ages, and this report will be based primarily on the scores of this constant sample of 33 cases.

These children had been coming in for mental and physical measurements from the time they were one month old. They had previously been given about 36 intelligence tests. With few exceptions[2] all tests were administered by the author. Starting at six years the tests were spaced at annual intervals, at or near the birth date. Tests of other, non-intellectual characteristics, were given at the six-month intervals between birthdays. It is evident, thus, that general practice effects had probably reached a saturation point long before the 16-year Wechsler-Bellevue was given. That is, the children knew me well, were familiar with the testing situation, and knew well the Stanford-Binet tests, which are similar to the Wechsler in procedure and types of items involved. There may, however, remain some specific memories from the first testing of the Wechsler that give clues for solution of some items at the later tests (2). Any such memories would have to carry over two- and three-year periods.

---

[2] Marjorie Honzik gave most of the two-year tests, and Mary Shirley the 10-year tests.

## C. RESULTS

The means and *SD*'s of scores earned by these 33 subjects are given in Table 1. These scores are based on the conventional 10 categories, although the vocabulary test was also given. The means are shown graphically in Figure 1. The top line represents the mean *IQ*. The group is superior, their *IQ*'s for the three ages averaging 119, 123, and 126; the *SD*'s are approximately 16 points. Theoretically, the mean *IQ* should not change, because Wechsler's tables of *IQ*s allow for some improvement over this age span. Taking into account the fact that this is a constant sample, and the scores between tests are highly correlated, this increase is statistically highly significant. It is, moreover, interesting to note that 30 of the 33 cases made *IQ* gains. The gains ranged from 1 to 20 points, while the greatest loss was only four points (see Figure 3). Some part of this increase in *IQ* may be due to specific practice effect. It is also possible that in using cross-section

TABLE 1

MEANS AND STANDARD DEVIATIONS OF WECHSLER-BELLEVUE INTELLIGENCE SCALE SCORES FOR 33 BERKELEY GROWTH STUDY CASES, TESTED AT 16, 18, AND 21 YEARS*

| Scale | 16 Years Mean | 16 Years *SD* | 18 Years Mean | 18 Years *SD* | 21 Years Mean | 21 Years *SD* |
|---|---|---|---|---|---|---|
| Full Scale *IQ* | 118.7 | 16.4 | 123.3 | 16.8 | 125.6 | 16.2 |
| Full Scale *EQ* | 115.5 | 14.7 | 121.4 | 15.7 | 125.6 | 16.2 |
| Performance *EQ* | 114.5 | 11.7 | 120.2 | 12.7 | 124.2 | 12.1 |
| Verbal *EQ* | 113.5 | 17.3 | 118.7 | 17.1 | 122.1 | 18.9 |
| *By Sexes Separately* | | | | | | |
| Full Scale *EQ* | | | | | | |
| 15 Boys | 114.7 | 17.4 | 121.3 | 19.2 | 124.9 | 20.0 |
| 18 Girls | 116.1 | 11.9 | 121.6 | 12.0 | 126.1 | 12.3 |
| Performance *EQ* | | | | | | |
| 15 Boys | 111.4 | 12.0 | 118.2 | 14.5 | 122.3 | 14.9 |
| 18 Girls | 117.2 | 10.7 | 121.8 | 9.7 | 125.7 | 8.9 |
| Verbal *EQ* | | | | | | |
| 15 Boys | 114.7 | 22.0 | 120.1 | 21.1 | 122.9 | 22.6 |
| 18 Girls | 112.4 | 11.8 | 117.4 | 12.8 | 121.4 | 15.0 |

*The five-year gains in these correlated scores are all significant above the .001 level of confidence.

populations Wechsler did not make adequate allowance for continued mental growth during this period.

The test-retest correlations in *IQ* are comfortably high (Table 2): for the full scale, the *r* between 16 and 18 years is .955, for 18 by 21 years years the *r* is .961, and for 16 by 21 years it is .951. For the 5-year span the Verbal Scale retest *r* is .90 and the Performance *r* is .85.

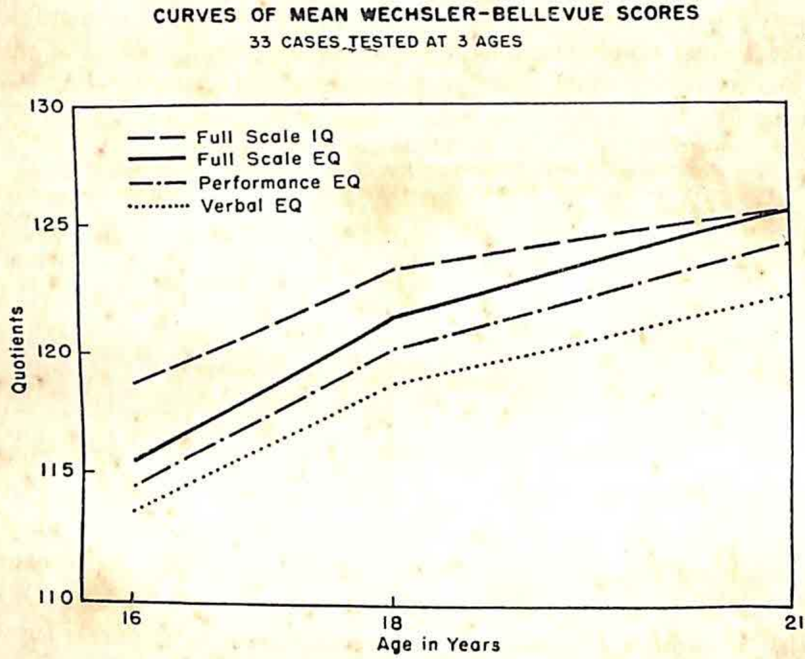

FIGURE 1
Age Curves of Wechsler-Bellevue Mean Scores for *IQ* and *EQ* for the Full Scale, and *EQ* for the Performance and Verbal Scales (Thirty-Three Cases of the Berkeley Growth Study)

The three more steeply-inclined lines in Figure 1 are means for Wechsler's Efficiency Quotients. Wechsler transposed the raw scores of each of the subtests into weighted scores according to their means and standard deviations for his total population. A person's Deviation *IQ* is read from tables based on the mean and *SD* of total weighted scores earned by people the same age. Wechsler found the highest scores occurring in the age-period 20 to 24 years. The Efficiency Quotient, therefore, is a score earned at any age, as expressed in terms of the 20-24-year norms. Changes in ability regardless of age may thus be expressed either in weighted score units or in *EQ*'s.

Accordingly, the *EQ* curves represent these children's progress toward (or deviations from) the 20-24-year norm. At 21 years, of course, the *IQ* and *EQ* are identical. In this sample all of the mean scores rise from 16 to 18 to 21 years, more steeply in the 16 to 18 interval. The 5-year increase in *EQ* is highly significant (Table 1), and appears to reflect true growth in ability. This is in line with the findings from cross-sectional studies.

Before we go on to analyze the characteristics of individual growth, we may note that, although the Performance Quotients are consistently higher than the Verbal Quotients, the Performance and Verbal halves of the scale show approximately equal gains, with trends that are almost parallel to the full scale.

TABLE 2

Correlation Coefficients for Wechsler-Bellevue Intelligence Scale Scores for 33 Berkeley Growth Study Cases Tested at 16, 18, and 21 Years

| Test | 2-Yr. Interval 16-18 | 3-Yr. Interval 18-21 | 5-Yr. Interval 16-21 |
|---|---|---|---|
| | *r* | *r* | *r* |
| Full Scale *IQ* | .955 | .961 | .951 |
| Verbal *IQ* | .902 | .924 | .895 |
| Performance *IQ* | .858 | .918 | .849 |

In Table 1 and Figure 2 the *EQ*'s are given according to sex. For these 15 boys and 18 girls, there seems to be no tendency for the girls to stop growing sooner than the boys: in this respect they evidently do not follow the physical sex differences in rates of maturing. It may be surprising to see that the Berkeley Growth Study boys do better on the verbal scale, while the girls are ahead in Performance. However, the same differences have been found in several other studies (5, 9).

Is gain in scores related to level of ability? The answer for this sample may be seen in Table 3 and Figures 3 and 4. Figure 3 is the scattergram of *IQ*'s at 16 years, in relation to the 5-year gain in *IQ*. The *r* is —.18: in other words, there appears to be no relation between 16-year *IQ* and subsequent intellectual growth. The scatter, however, has an interesting form. Those with moderate or low *IQ*'s seem to have a steady gain: among those with *IQ*'s over 100, we find both the extreme gains and the few losses in *IQ*.

In Figure 4, changes are expressed in units of intellectual growth instead of quotients, and as progress toward a "terminal" status: that is, five-year gains in weighted scores are shown in relation to the 21-year *IQ*'s. The *r* is .45, though it is not linear. If we correlate these same 5-year gains with the 16-year *IQ*'s the *r* drops to .26, which is not significant. This way of showing units of growth in intelligence in terms of weighted scores also gives evidence of steady gain in those who have low *IQ*'s. And here, too, the more intelligent subjects seem to be in varying stages toward reaching their top capacity. Perhaps the one girl who showed no gain had reached her intellectual ceiling by 16 years, and several others were approximately there. Still others continued to make rapid strides in tested ability.

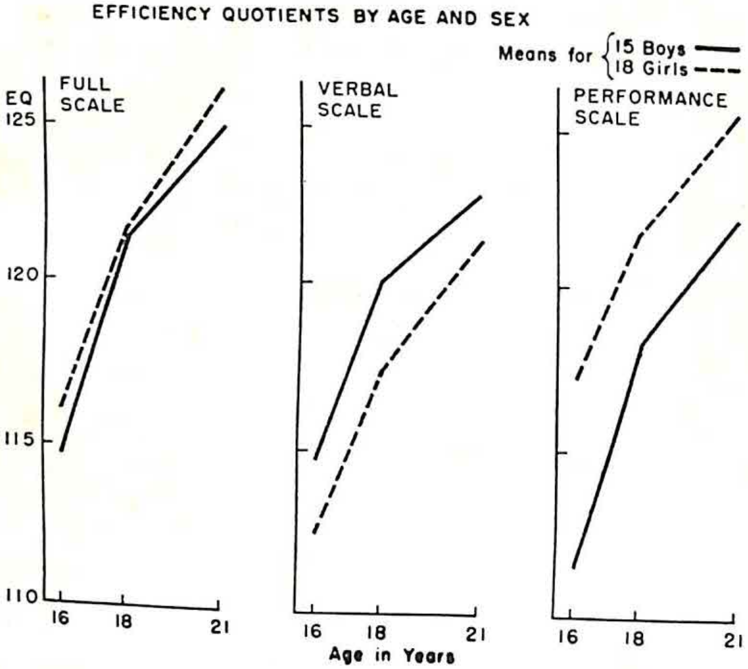

FIGURE 2

AGE CURVES OF WECHSLER-BELLEVUE MEAN *EQ* SCORES ACCORDING TO SEX, FOR THE FULL SCALE AND FOR THE VERBAL AND PERFORMANCE SCALES

What about education? These correlations are given in Table 3. The 16-year *EQ* correlates .84 with years of schooling completed by 21 years,

TABLE 3

CORRELATIONS BETWEEN SCORES AND LEVEL OF EDUCATION

| Scores | $r$ |
|---|---|
| 21-Year *IQ* (or *EQ*) | .82 |
| 21-Year *EQ* | .84 |
| 5-Year Gain in *EQ* | .33* |

*$P$ is .06.

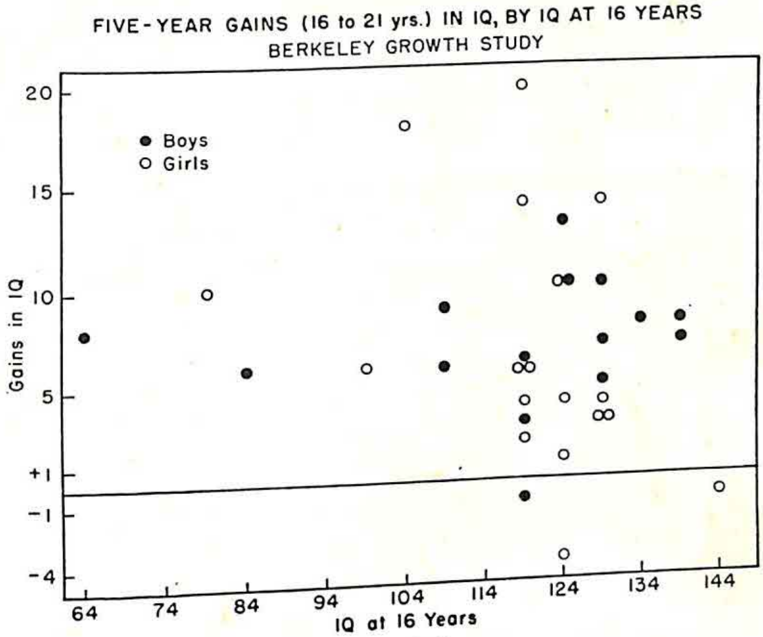

FIGURE 3
SCATTERGRAM OF FIVE-YEAR GAINS IN *IQ* BY INITIAL *IQ* AT 16 YEARS

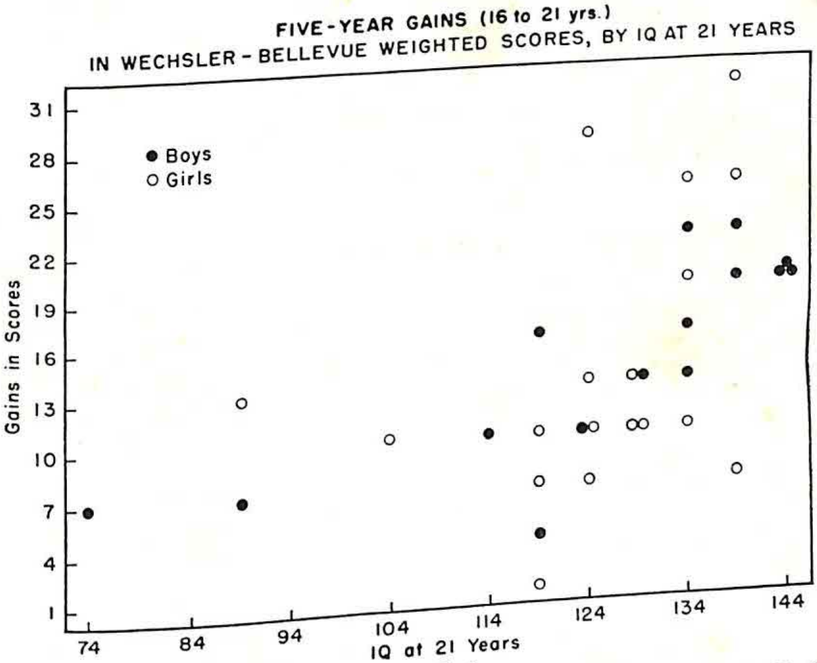

FIGURE 4
SCATTERGRAM OF FIVE-YEAR GAINS IN WEIGHTED SCORES BY FINAL *IQ* AT 21 YEARS

while the 21-year quotient correlates .82 with years of schooling. The 5-year *gain* in *EQ* correlates .33 with years of schooling, an *r* that misses being significant for the small sample: the level of confidence is .06.

The relative independence of gain and education may be seen in some individual cases. Figure 5 shows five pairs of cases, selected for closely similar *EQ*'s at 16 years. The broken line in each pair represents the member with higher education. In all five pairs the case with higher schooling went beyond the University A.B. degree. Three are medical students, one is in law, and one boy has earned (after the age of 21 years) an M.S. in one of the physical sciences. The cases with lower education are only moderately low: they run from 3 years high school to the A.B. degree in a University. The patterns of growth are about as varied as five pairs could be: they do not show any strong evidence that continuing education is the determining factor. It is interesting to note, however, that the two girls who are included here both drop in *EQ* at 18 years. Both were married at about this age, and the high school girl, whose scores do not increase, had children be-

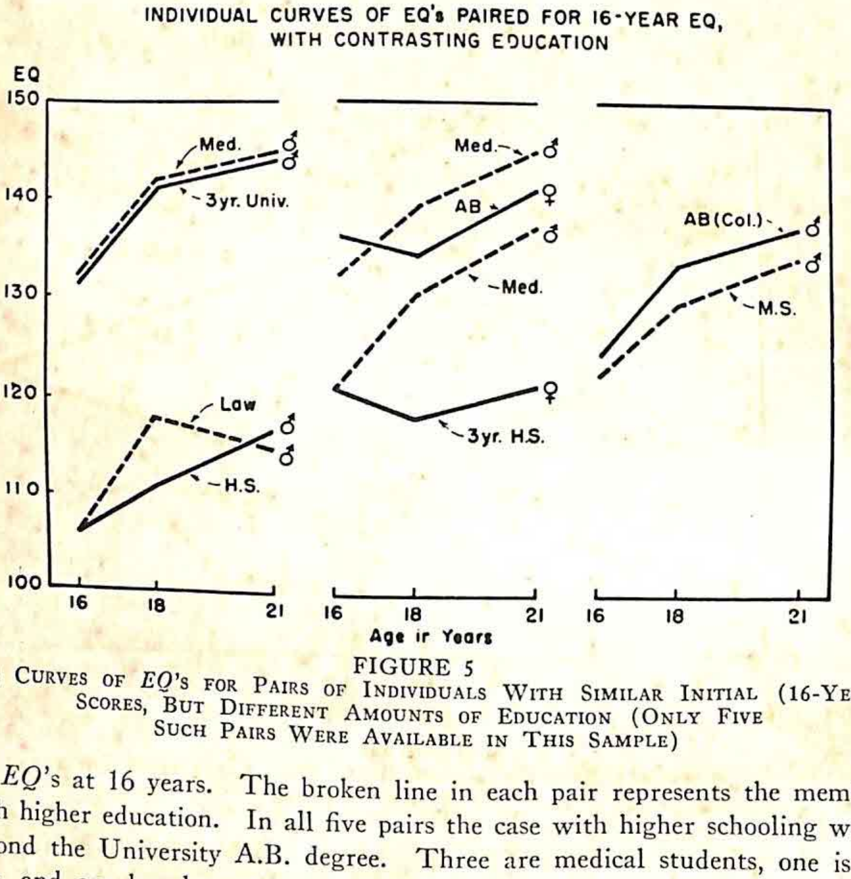

FIGURE 5
AGE CURVES OF *EQ*'S FOR PAIRS OF INDIVIDUALS WITH SIMILAR INITIAL (16-YEAR) SCORES, BUT DIFFERENT AMOUNTS OF EDUCATION (ONLY FIVE SUCH PAIRS WERE AVAILABLE IN THIS SAMPLE)

fore she was 21. In the other three pairs, all boys, the member with higher education did not seem to have any advantage.

In Figure 6 is a pair of identical twin girls. They went through high school together, with very similar life experiences. At 16 Twin *A* scored seven *EQ* points higher than Twin *B*. Between 16 and 18, Twin *B* was

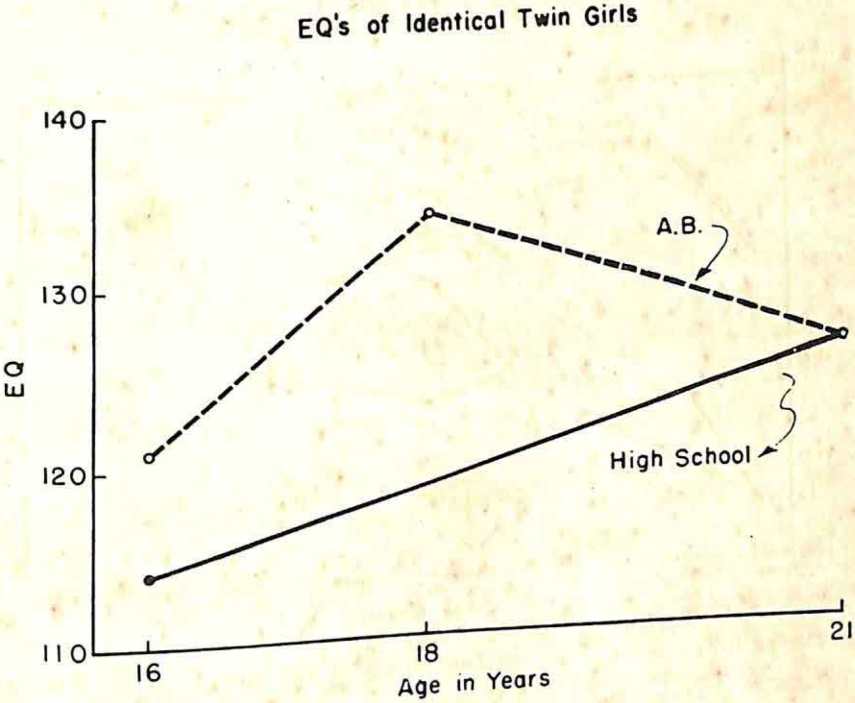

FIGURE 6
AGE CURVES OF *EQ*'S FOR A PAIR OF IDENTICAL TWIN GIRLS WITH DIFFERENT AMOUNTS OF EDUCATION

married, moved to another part of the country, and had no more education. Twin *A* went to the University and in the course of time graduated. *B* was not available for testing at 18 years, but at 21 years the girls made identical scores.

There were three children who dropped out of school because of obvious lack of ability. Their *IQ*'s at 16 years were 64, 75, and 84. The boy with the lowest score never learned to read, and remained in an ungraded class. The other boy and girl were for two years in "continuation school," i.e., worked part time and went to high school one or two days a week until they were 18. All three gained moderately, after 16 years, from 6 to 10 *IQ* points (see Figure 3).

These comparisons of cases are only suggestive. We cannot, in our small sample, effectively separate the factors of education and level of intelligence as they are related to mental growth. But it is evident that growth did occur in all levels of education and *IQ* that are represented.

TABLE 4
RELATION OF ABILITY-LEVEL TO VERBAL/PERFORMANCE RATIO

| Full scale *EQ* with ratio at | *r* |
|---|---|
| 16 Years | .48 |
| 18 Years | .46 |
| 21 Years | .62 |

TABLE 5
CORRELATIONS BETWEEN INTELLIGENCE LEVEL AND GAIN IN SCORES

| Items correlated | Age intervals 16-18 *r* | 18-21 *r* | 16-21 *r* |
|---|---|---|---|
| *Full Scale:* | | | |
| 21-Year *IQ* with gain in *IQ* | | | .13 |
| 21-Year *IQ* with gain in *EQ* | .23 | .27 | .45* |
| 16-Year *IQ* with gain in *IQ* | | | —.18 |
| 16-Year *IQ* with gain in weighted score | | | .26 |
| *Verbal Scale:* | | | |
| 21-Year *IQ* with gain in *EQ* | | | .42** |
| 16-Year *EQ* with gain in *EQ* | | | .02 |
| *Performance Scale:* | | | |
| 21-Year *IQ* with gain in *EQ* | | | .33 |
| 16-Year *EQ* with gain in *EQ* | | | —.21 |

*Significant at .01 level of confidence.
**Significant at .02 level of confidence.

Thus far we have talked mostly of differences in the growth of individuals. But there is the other problem of differential growth rates found in different types of mental function. We have seen in Table 1 and Figure 1, that both the Performance and Verbal scores increased during this five-year period. Although the Performance scores average consistently higher than the Verbal, those subjects with the higher *IQ*'s did relatively better in the Verbal Scale. This may be seen in Table 4 by the correlations between *EQ*'s and the Verbal Performance ratios.[3] The *r* is .48 at 16 years, .46 at 18, and .62 at 21 years. On the other hand, as may be seen in Table 5, the correlation between the Verbal *EQ* at 16 years and the five-year gain in Verbal scores is only .02: the similar comparison in Performance scores yields an *r* of —.21. That is, although the brighter subjects in this group are relatively better in the verbal tests, there is no tendency for them to make greater gains in these tests.

---

[3] Verbal Weighted Score divided by Performance Weighted Score.

In Table 6 and Figure 7 are presented the data for weighted scores for each of the 11 separate categories. These means are based on the total sample tested, but are very similar to those obtained for the constant sample of 33 cases.

TABLE 6

MEANS AND *SD*'S OF WECHSLER-BELLEVUE WEIGHTED CATEGORY SCORES FOR TOTAL BERKELEY GROWTH STUDY SAMPLE

| | 16 Years (*N* is 45) | | 18 Years (*N* is 40) | | 21 Years (*N* is 35) | |
|---|---|---|---|---|---|---|
| Category | Mean | *SD* | Mean | *SD* | Mean | *SD* |
| Information | 11.42 | 2.82 | 12.08 | 2.72 | 12.43 | 3.06 |
| Comprehension | 13.07 | 2.84 | 14.25 | 2.70 | 14.43 | 3.17 |
| Digit Span | 10.20 | 3.11 | 10.92 | 3.14 | 11.66 | 3.48 |
| Arithmetic | 10.58 | 3.57 | 11.22 | 3.95 | 11.89 | 4.17 |
| Similarities | 12.31 | 3.18 | 13.28 | 3.03 | 13.66 | 3.38 |
| Vocabulary | 11.54 | 2.04 | 11.95 | 2.17 | 12.43 | 2.44 |
| Picture Arrangement | 11.13 | 2.86 | 11.58 | 3.40 | 12.49 | 3.09 |
| Picture Completion | 12.47 | 2.25 | 13.28 | 1.63 | 13.20 | 1.24 |
| Block Design | 13.34 | 2.44 | 14.08 | 2.24 | 14.80 | 2.42 |
| Digit Symbol | 11.43 | 2.19 | 12.50 | 2.45 | 13.17 | 2.47 |
| Object Assembly | 11.71 | 2.73 | 13.44 | 1.77 | 13.95 | 1.80 |

When the scores are broken down into categories several interesting things come out. It is clear that for our sample the different categories are not of equal difficulty (5, ). Although all categories show an increase in scores, in some the change is much greater than in others.

In some categories, such as Picture Completion and Comprehension the group seems to have reached or approached a ceiling. In other categories the scores appear to be still advancing. Picture Arrangement, a relatively difficult test, seems actually to accelerate between 18 and 21 years. These differences could be due to differences in the growth rates of the functions tested, to differential memory for previous solutions of the problems, to difference in the adequacy with which the tests measure the various functions, or to all of these causes. In several instances, for example, it seems likely that if more difficult items were added to a category, we might find more continuous growth in it.

The Vocabulary Scale is included here, although it was not used in computing the *EQ*'s or *IQ*'s. It is interesting to note that this test is intermediate in its difficulty and in its general tendency to show continuous growth over the entire five years.

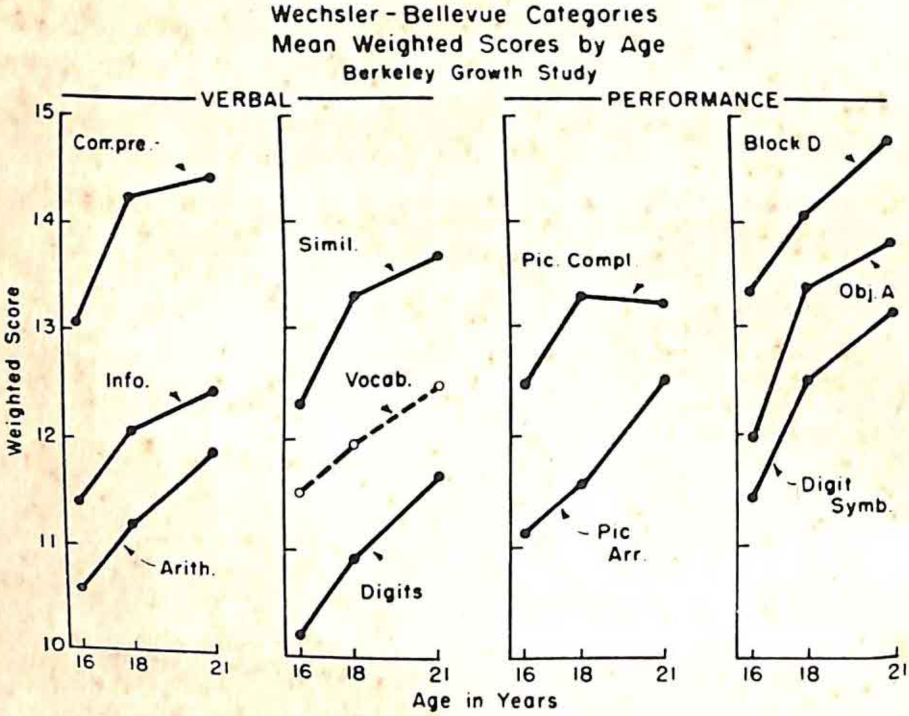

FIGURE 7
AGE CURVES OF THE MEANS OF WECHSLER-BELLEVUE CATEGORY SCORES FOR THE TOTAL BERKELEY GROWTH STUDY SAMPLE

## D. SUMMARY

In this small group of 33 cases we have evidence that intelligence, as measured by repeated tests on the Wechsler-Bellevue scale, increases (about equally for both sexes) from 16 to 18 to 21 years of age. The gains occur at all levels of intelligence and education found in this sample. There are indications that some persons appear to have reached their own top capacities by 16 to 18 years, while others may still be growing at 21.

As for the kinds of ability tested in the Wechsler Scale, these subjects (whose *IQ*'s average high) did better on the Performance Scale, though in general the higher the *IQ* the relatively better the Verbal Scale score. When the scale is broken down into the separate categories, the categories were found to be unequal in difficulty for this group. All categories showed some increase in scores with age, but some appeared to reach their ceiling by 18 years and others were still showing increases at 21 years.

These findings indicate that intellectual functions continue to change through 21 years and probably beyond. Tests on this sample at 25 years, now in progress, should yield further information about the developmental trends of intelligence in young adults.

## REFERENCES

1. Bayley, N. Consistency and variability in the growth of intelligence from birth to eighteen years. *J. Genet. Psychol.*, 1949, **75**, 165-196.
2. Derner, G. F., Aborn, M., & Canter, A. H. The reliability of the Wechsler-Bellevue sub-tests and scales. *J. Consult. Psychol.*, 1950, **14**, 172-179.
3. Foulds, G. A. Variations in the intellectual activities of adults. *Amer. J. Psychol.*, 1949, **62**, 238-246.
4. Freeman, F. N., & Florey, C. D. Growth in intellectual ability as measured by repeated tests. *Monog. Soc. Res. Child. Devel.*, 1937, **2**, 1-116.
5. Jastak, J. Ranking Bellevue subtest scores for diagnostic purposes. *J. Consult. Psychol.*, 1953, **17**, 403-410.
6. Jones, H. E., & Bayley, N. The Berkeley Growth Study. *Child Devel.*, 1941, **12**, 167-173.
7. Jones, H. E., & Conrad, H. S. The growth and decline of intelligence: a study of a homogeneous group between the ages of ten and sixty. *Genet. Psychol. Monog.*, 1933, **8**, 233-297.
8. Miles, W. R. Psychological aspects of ageing. In *Problems of Ageing* (Sec. Ed.). Baltimore: Williams & Wilkins, 1942. (Pp. 756-84.)
9. Norman, R. D. Sex differences and other aspects of young superior adult performance on the Wechsler-Bellevue. *J. Consult. Psychol.*, 1953, **17**, 411-418.
10. Raven, J. C. The comparative assessment of intellectual ability. *Brit. J. Psychol.*, 1948, **39**, 12-19.
11. Thorndike, R. L. Growth of intelligence during adolescence. *J. Genet. Psychol.*, 1948, **72**, 11-15.
12. Wechsler, D. The Measurement of Adult Intelligence. New York: Williams & Wilkins, 1944. Pp. 258.

*Child Development Section*
*National Institute of Mental Health*
*Bethesda 14, Maryland*

*The Journal of Genetic Psychology*, 1957, **90**, 17-22.

# A COMPARATIVE STUDY OF THE INTELLIGENCE OF JEWISH PRESCHOOL BOYS AND GIRLS OF ORTHODOX PARENTAGE*

*Department of Psychology, Yeshiva University*

Boris M. Levinson[1]

## A. Introduction

This study is part of a larger research of the intelligence, personality, interests, and motor abilities of Jewish children attending day (yeshiva) schools. The purpose of this study was to learn whether there were any statistically significant differences in the intelligence of Jewish preschool boys and girls attending day schools. It was predicted that boys would have a statistically significantly higher *IQ* because of the differential motivation and cultural pressure exerted upon them as compared with the girls. It is a well known fact that among Jewish Orthodox parents, the boy is looked upon as the one who will transmit the traditions of the family, nation, and religion. It is the boy who is under religious obligation to study; who is expected to accomplish more in his studies than his sister and is given preference when it comes to a choice as to which member of the family is to attend day school. The father pays almost exclusive attention to him to the relative neglect of the girl. There is thus created a very strong bond of affinity between the father and son on one side and daughter and mother on the other. It was felt that the varying rôles, social and religious status of the boy and girl would tend toward the development of different attitudes with resultant differing social selves and personalities. It was thus expected that the cultural imperative for achievement on the part of the boys would lead toward a statistically significant difference in the intelligence of these preschool children in favor of the boys.

## B. Previous Research

A review of literature pertaining to this problem does not reveal any investigation in the exact area of this study. There are, however, some

---

*Received in the Editorial Office on July 14, 1955.

[1] The writer wishes to express his appreciation to Professor Joshua B. Matz, who supervised and checked the statistical work, and to Morton Schwartzstein and Richard Garber, who tabulated and computed the statistical data.

related studies. These may be divided into (*a*) studies of the intelligence of Jewish children and (*b*) the effect of cultural pressure upon children's intellectual development. Critical summaries of previous studies of the intelligence of Jewish children were made by Maller (11), Brill (2), and Nardi (13). Davies and Hughes (4) made a study of 1,894 school children ages eight to 14 to whom the Northumberland Standardized Tests were administered. They found that in intelligence the Jewish boys were slightly superior to Jewish girls. The reverse was true among non-Jewish children. Winch (19) found that while Jewish children attending East End elementary schools in London were superior intellectually to gentile boys, the difference between Jewish and gentile boys was greater than between Jewish and gentile girls. Brown (3) made a study of second generation Jewish children attending public school kindergartens: 185 boys and 150 girls were examined on the 1916 Revision of the Stanford-Binet Scale. The mean *IQ* of the boys was 107.5 and of the girls 109.3 (as computed by writer from data given by Brown). The difference in *IQ* is not statistically significant at the .05 level. Brown (3) noted that "the Jewish child is more critical of himself as a performer and less critical of his objective performance."

It is to be noted that the children who were subjects of these investigations were attending public schools and possibly the cultural bias was not felt by them as sharply. Furthermore, the problem of differential cultural pressure in the Jewish home on the intellectual development of boys and girls were not considered by these investigators.

Research studies in the general area of intellectual growth have indicated that the child's mental development is to a large extent dependent upon the socio-economic level of his home (3, 16) and that the child's immediate home environment can serve either to develop or retard certain intellectual potentialities (14, 15). Thus at 10, children from upper socio-economic groups are superior to children from lower socio-economic groups in both verbal and performance areas (6). At 16, they are better only in verbal areas (7). Sex differences in intellectual ability are also more pronounced among children of low socio-economic background (5, 20). The effect of cultural pressure may further be seen in the fact that whereas there is an "excess of boys over girls in higher *IQ* brackets in the majority of groups of gifted generally" (12), this is reversed among Negroes where there are more high ranking girls than boys (8). We also find that contrary to the general trend some Negro preschool boys speak longer sentences than do Negro preschool girls (1). McCarthy (9) feels that this may be due to the different attitudes toward the two sexes which may bring about different treatment and hence different performances.

## C. Procedure

One thousand cases of Jewish preschool children examined at the Educational Advisement Center of Yeshiva University were selected for study. These were normal children who were referred for intelligence, personality, and school readiness examinations prior to placement in first grade of yeshiva schools. Some of these children were examined while attending the kindergarten. The intelligence scale administered to these children was the Revised Stanford-Binet, Form *L* (16). The preschool age chosen for study was from 4 years 3 months to 6 years 5 months. It was felt that the Revised Stanford-Binet was a fair measure of the intelligence of these children at this age since it does not favor either sex. The writer computed the mean *IQ*'s for both boys and girls (ages 4½ to 6) on the standardization sample of Revised Stanford-Binet. Using the data supplied by McNemar (10, p. 44) he found that the mean *IQ* (*N* of 262) of the boys was 102.71, and that of the girls of the same age (*N* of 260) was 102.84. This difference is not statistically significant. The observation made that girls usually do better than boys on verbal tests, which may or may not be valid, does not hold true for the Revised Stanford-Binet at the age level under consideration. Items in which girls are superior on the preschool level involve "some type of manipulative ability or discrimination or number ability" (10). Girls who come from a bilingual home are thus not penalized as against boys who come from the same background.

Of the one thousand children selected for the study, 770 were between the ages of 4 years 3 months and 6 years 5 months. Of these 484 were boys and 286 were girls (see Table 1). This group will be referred to as Sample *A*. Apparently some selective factors were affecting this group inasmuch as more boys than girls were candidates for admission to the day schools. To eliminate as far as possible all extraneous factors, outside of sex, all male and female siblings or a total of 89 children, 45 boys and 44 girls were selected (see Table 2). This group will be referred to as Sample *B*. Sample *C* consisted of the remaining group of 439 boys and 242 girls.

## D. Results

Table 1 shows the distribution of the *IQ*'s and *CA*'s of Sample *A*. The mean *IQ* of the boys is 112.8 with a *SD* of 12.74; the mean *IQ* of the girls is 113.6 with a *SD* of 13.56. The difference between the means is statistically not significant.

Table 2 shows the distribution of the *IQ*'s of Sample *B*. The mean *IQ* of the boys is 118.06 with a *SD* of 14.0; the mean *IQ* of the girls is 114.2

TABLE 1
SEX DIFFERENCES IN *IQ*'S OF JEWISH PRESCHOOL CHILDREN

| Age | *N* | Sex | *IQ* | *SD* |
|---|---|---|---|---|
| 4:3-4:5 | 10 | Boys | 127.3 | 16.48 |
| | 5 | Girls | 110.4 | 5.95 |
| 4:6-4:8 | 18 | Boys | 114.27 | 12.78 |
| | 9 | Girls | 121.22 | 19.86 |
| 4:9-4:11 | 20 | Boys | 116.45 | 14.58 |
| | 11 | Girls | 118.09 | 13.80 |
| 5:0-5:2 | 43 | Boys | 119.11 | 11.02 |
| | 25 | Girls | 119.68 | 11.00 |
| 5:3-5:5 | 97 | Boys | 113.22 | 12.37 |
| | 53 | Girls | 114.86 | 13.80 |
| 5:6-5:8 | 94 | Boys | 112.48 | 13.58 |
| | 59 | Girls | 113.42 | 13.14 |
| 5:9-5:11 | 98 | Boys | 109.11 | 12.63 |
| | 64 | Girls | 112.50 | 12.88 |
| 6:0-6:2 | 67 | Boys | 109.28 | 11.98 |
| | 42 | Girls | 109.00 | 14.74 |
| 6:3-6:5 | 37 | Boys | 114.83 | 15.91 |
| | 18 | Girls | 111.61 | 11.80 |

TABLE 2
SEX DIFFERENCES IN SIBLING *IQ*'S OF JEWISH PRESCHOOL CHILDREN

| *IQ* | Boys | Girls |
|---|---|---|
| 70- 79 | — | 1 |
| 80- 89 | 2 | 2 |
| 90- 99 | 3 | 4 |
| 100-109 | 6 | 10 |
| 110-119 | 9 | 11 |
| 120-129 | 18 | 10 |
| 130-139 | 5 | 2 |
| 140-149 | 2 | 3 |
| 150-159 | — | 1 |

with a *SD* of 16.90. The difference between the means is statistically not significant.

In Sample *C* the boys had an *IQ* of 112.11 with a *SD* of 12.38 the girls had an *IQ* of 113.11 with a *SD* of 12.84. The difference between the means is statistically not significant.

A comparison between the mean *IQ*'s of the boys in Samples *B* and *C* indicates, however, that there is a statistically significant difference at the .05

level of confidence. There is no statistically significant difference between the mean *IQ*'s of the girls in Samples *B* and *C*.

## E. DISCUSSION

While it is true that the slight superiority in intelligence of the Jewish preschool Orthodox boy as compared with the girl is statistically not significant, it does indicate a trend which merits further study. It is to be remembered that our sample is a small one and, therefore, of necessity differences in means unless very large would not be statistically significant.

This difference in *IQ* may also indicate the beginning of personality differences in areas other than intellectual which are not measurable by the Revised Stanford-Binet. It is to be observed that apparently even at the preschool level the Jewish child is already under pressure to achieve. In this connection note Terman and Oden's statement (17): "The Jewish child is under heavy pressure to succeed, with the result that he achieves more per unit of intelligence than those of other racial stocks."

It is interesting to note that the girls' mean *IQ*'s in Samples *B* and *C* are almost identical and hence most likely represent the same population and possibly the same socio-economic level. There is, however, a statistically significant difference in the *IQ*'s of boys of Samples *B* and *C*. Apparently the majority of the boys may have come either from homes on a lower socioeconomic plane with parents financially unable to send both sexes to day schools or from families where it is not considered necessary for a girl to have a day school education.

## F. SUMMARY AND CONCLUSIONS

A study of the intelligence of Orthodox Jewish preschool male (*N* of 45) and female (*N* of 44) siblings indicated that while the *IQ*'s of the boys (118.06) was higher than that of the girls (114.2), the difference was not statistically significant. This finding was contrary to the hypothesis that differential cultural pressure at home would lead to a significantly higher *IQ* on part of the boys as compared with the girls. It was felt, however, that since this sample represented the total population out of a thousand children, a possible trend may be indicated which should be studied further. It is also possible that while cultural pressure for high achievement may have a somewhat equally stimulating effect on intellectual abilities in both sexes, the cultural imperative may differentially modify the personality of the preschool children in areas not measurable by the Stanford-Binet.

J-312
3.9.70

## References

1. Anastasi, A., & D'Angelo, R. A comparison of Negro and white preschool children in language development and Goodenough draw-a-man *IQ*. *J. Genet. Psychol.*, 1952, **81**, 147-165.
2. Brill, M. Studies of Jewish and non-Jewish intelligence. *J. Educ. Psychol.*, 1936, **27**, 331-352.
3. Brown, F. A comparative study of the intelligence of Jewish and Scandinavian kindergarten children. *J. Genet. Psychol.*, 1944, **64**, 67-92.
4. Davies, M., & Hughes, A. G. An investigation into the comparative intelligence and attainments of Jewish and non-Jewish children. *Brit. J. Psychol.*, 1927, **18**, 134-146.
5. Davis, E. A. The development of linguistic skill in twins, singletons with siblings, and only children from age five to ten years. *Inst. Child Welfare Monogr. Ser.*, No. 14. Minneapolis: Univ. Minnesota Press, 1937.
6. Havighurst, R. J., & Janke, L. L. Relations between ability and social status in a mid-Western community: I. Ten-year-old children. *J. Educ. Psychol.*, 1944, **35**, 357-369.
7. Janke, L. L., & Havighurst, R. J. Relations between ability and social status in a mid-Western community: II. Sixteen-year-old boys and girls. *J. Educ. Psychol.*, 1945, **36**, 499-509.
8. Jenkins, M. D. A socio-psychological study of Negro children of superior intelligence. *J. Negro Educ.*, 1936, **5**, 175-190.
9. McCarthy, D. Some possible explanations of sex differences in language development and disorders. *J. of Psychol.*, 1953, **35**, 155-160.
10. McNemar, Q. The Revision of the Stanford-Binet Scale. New York: Houghton, 1942.
11. Maller, J. B. Studies in the intelligence of young Jews. *Jewish Educ.*, 1931, **2**, 29-39.
12. Miles, C. C. Gifted children. In (L. Carmichael, *Ed.*), *Manual of Child Psychology* (2nd Ed.). New York: Wiley, 1954. (Pp. 984-1063.)
13. Nardi, N. Studies in intelligence of Jewish children. *Jewish Educ.*, 1948, **19**, 41-50.
14. Skeels, H. M., & Harms, I. Children with inferior social histories: Their mental development in adoptive homes. *J. Genet. Psychol.*, 1948, **72**, 283-294.
15. Skodak, M. Mental growth of adopted children in the same family. *J. Genet. Psychol.*, 1950, **77**, 3-9.
16. Terman, L. M., & Merrill, M. A. Measuring Intelligence. New York: Houghton, 1937.
17. Terman, L. M., & Oden, M. Status of the California Gifted Group at the end of sixteen years. Thirty-Ninth Yearbook National Society for the Study of Education. *Intelligence, Its Nature and Nurture*: Part 1, 1940.
18. ———. The Gifted Child Grows Up. Stanford University: Stanford Univ. Press, 1947.
19. Winch, W. H. Christian and Jewish children of East End elementary schools. *Brit. J. Psychol.*, 1930, **20** (Pt3), 261-273.
20. Young, F. M. An analysis of certain variables in a developmental study of language. *Genet. Psychol. Monog.*, 1941, **23**, 3-141.

*Department of Psychology*
*Yeshiva University*
*506 West 185th Street*
*New York City 33*

*The Journal of Genetic Psychology*, 1957, **90**, 23-32.

# PROBABILITY LEARNING IN CHILDREN: SOME EXPLORATORY STUDIES*

*Department of Psychology, University of Illinois*

SAMUEL J. MESSICK AND CHARLES M. SOLLEY

## A. INTRODUCTION

Brunswik (1, 2) has suggested that learning processes should be studied under conditions more nearly approximating everyday life situations. He pointed out that the classic experimental procedure of rewarding responses to one stimulus all of the time while never rewarding responses to another stimulus is unrealistic. In everyday life situations a specific response to a specific stimulus class has a probability generally less than one of being rewarded or "correct." This formulation emphasizes the weighting of response decisions in terms of previously learned probabilities. For example, if a stimulus event has a probability of .75 of being followed by reinforcement, then *S* might begin responding to that stimulus more and more frequently until he is choosing it 75 per cent of the time. Or he might respond to it all of the time in order to maximize the number of reinforcements obtained.

The probability of a stimulus being rewarded if chosen all of the time is called *stimulus input probability* or, more simply, just *input probability*. The proportion of times *S* chooses or responds to that stimulus is commonly called *response output probability* or, equivalently, just *output probability*. Depending upon the purpose of the experiment and on the complexity of the stimulus choices, the experimenter may utilize only these probability definitions, or he may choose to work with expectation values of a log transformation of the probabilities, usually called "bits of information." The present authors propose to use only the probability definitions in this series of studies.

The systematic change in output probability as a function of having more and more experience with an input probability is called *probability learning*. Studies in this area have used either adult college students (3, 6, 8, 9) or subhuman species such as rats (1) or goldfish (4). Differences obtained between studies could well be due to (*a*) different mathematical models being required for different experimental designs, (*b*) instructions given *S*s,

---

*Received in the Editorial Office on July 18, 1955.

(*c*) motivational variables, or (*d*) developmental differences between or within species. Human maturational and learning differences at different age levels could be investigated by using children as *S*s in these probability learning experiments, and the present studies attempt such an investigation of developmental differences at an exploratory level. These studies were designed to explore children's behavior in situations similar to those used with adults. Rather than use a large number of children in an extremely restricted set of conditions, it was decided to use a small number of children in a large number of different conditions for probability learning, with long enough time delays between conditions to avoid correlation of error components.

One purpose of the present experiments was to see whether or not young children would tend to "maximize probability" of reward in their guessing behavior. For example, if one stimulus is correct 75 per cent of the time and a second, alternative stimulus is correct the other 25 per cent of the time, then *S* will usually obtain as much reward as possible in the situation if he chooses the most frequently correct stimulus all of the time. *Maximizing probability*, then, means that *S* chooses the most frequently correct stimulus all of the time. If the young children tend to maximize probability of reward, then their output probability curves will converge toward a probability of 1.00 even though the input probability level is much lower.

A second question asked in these experiments was "How near chance probability (.50 in the two-choice situation used) could children discriminate a deviation from chance probability?" Brunswik (1) suggests that this "probability threshold" for rats is near .70. Data for humans (5) indicate that adults can at least discriminate a probability of .62 from .50 in a two-choice situation. In the present study the input values of 1.00, .90, .75, and .60 were used to bracket roughly the "probability threshold" for children. It was felt that the discrimination threshold for children would probably be higher than for adults and, hence, that they would not discriminate an input of .60 from chance. It was also felt that thresholds might vary as a function of the children's chronological age.

Three studies were done in an attempt to answer these questions. In the first study the questions concerning "maximizing probability" and "probability thresholds" were investigated. In the second study a stimulus with "social" characteristics was used—stick-men drawings with happy and sad faces. It was felt that perhaps the *S*s would split into two or more groups with respect to initial probability bias and rate of learning, with the curves of both groups increasing to an asymptote at the input probability. This notion was based upon an expectation that there would be more variation in the frequency with which such characteristics as "happy" and "sad" are experienced among children than for such physical characteristics as "big" and "little." In the third study the effect of adding an incentive to the reinforcement word "right" was investigated, i.e., in addition to telling *S* "right" he was given a small piece of candy as well. Experiment III was set up on the basis of the results of Experiment I. It was felt that with increased incentive the children would "maximize probability" of getting a reward.

## B. METHOD

### 1. *Subjects*

Subjects consisted of one eight-year-old child, one seven-year-old, three five-year-olds, one four-year-old, and one three-year-old. All of these children were judged to be above average in intelligence although no *IQ* test was given, and they were all normal and healthy.

### 2. *Materials*

*a. Experiment I.* The materials consisted of 4 x 6 index cards, upon which drawings of kangaroos, horses, and stick-men were mounted. In this study the figures varied only with respect to size (big-little). Decks of 20 cards were arranged, and samples of "big" and "little" stimuli were shown to the children before the learning task. One deck (horses) had an input probability of 1.00 for "big" and .00 for "little." Another deck (stick-men) had an input probability of .90 for "big" and .10 for "little." A third deck (also stick-men) had an input probability of .75 for "big" and .25 for "little." A fourth deck (kangaroos) had inputs of .60 and .40 for "big" and "little," respectively. Different kinds of stimuli—horses, stick-men, and kangaroos—were used to keep up the children's interest and to make the tasks appear different.

*b. Experiment II.* The only stimuli in this study were stick-men with happy and sad faces. The characteristic "happy"' had an input probability of .75, while "sad" had an input probability of .25.

*c. Experiment III.* In this study the stimuli were the same stick-men used in Exp. I for the .75 condition. The characteristic "big" occurred 75 per cent of the time and "little" occurred the other 25 per cent of the time, thus reproducing exactly the input conditions for the .75 level of Exp. I. However, in addition to telling *S* "right" when he guessed the card correctly, he was also given a small piece of candy.

### 3. *Procedure*

All of the children were tested separately, usually in their own homes. The children were told that they were about to play a guessing game. The procedure was essentially the same for all the three experiments. Before each experiment *S* was shown an example of each of the two figures, e.g., a "big" kangaroo and a "little" kangaroo. He was then asked to guess which figure would be on top of the deck. *S* was instructed to guess either "big" or "little" in Exps. I and III and to say either "happy" or "sad" in Exp. II. The deck of cards was then shuffled and placed face down on the table.

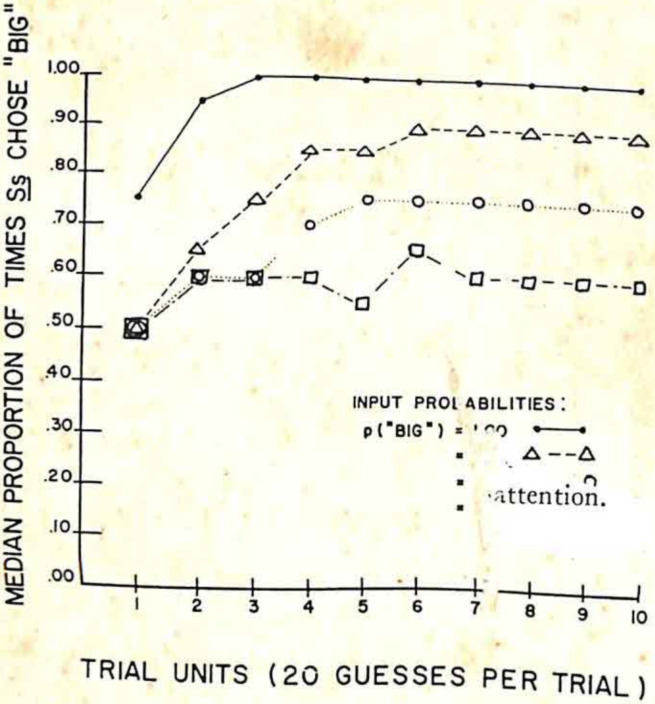

FIGURE 1

Probability Learning Curves for Four Stimulus Input Levels: All Ages Pooled, $N = 7$ for Each Curve

After *S* made each guess, the top card was turned over for his inspection. *S* was told "right" if he had guessed it correctly and "wrong" if he had guessed it incorrectly. In Exp. III *S* was also given a piece of candy if he was correct and nothing if he was wrong. When all 20 cards in the deck had been guessed, the deck was shuffled and the procedure repeated until *S* had gone through the deck 10 times for a total of 200 guesses. These

200 guesses were then arbitrarily divided into 10 trial units of 20 guesses each. The proportion of times *S* guessed "big" out of the 20 guesses in each trial unit was taken as his response probability for that trial (in Exp. II the measure was for "happy").

Rest periods were permitted whenever *S*s appeared to be tiring. Frequently the children would stop spontaneously, talk, walk around, get a drink of water, etc., before completing the task. The three- and four-year-old children stopped most frequently, were the most distractible, and often seemed to pay no attention whatever. Each series of 200 guesses was completed in a single session, and at least one day and not more than three weeks elapsed between sessions.

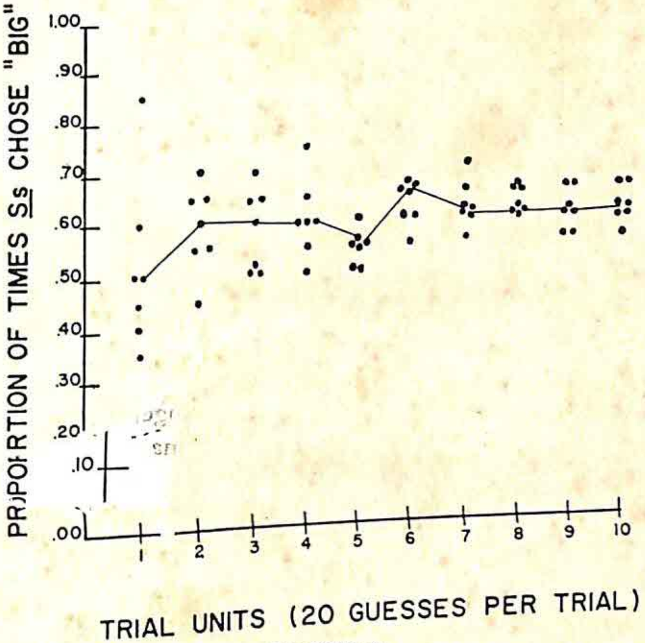

FIGURE 2
CHANGE IN SCORE DISTRIBUTION AT .60 INPUT LEVEL FOR SEVEN SUBJECTS

## C. RESULTS

### 1. *Experiment I*

The results of Exp. I are summarized in Figures 1 and 2. Figure 1 shows that the median response probabilities reach an asymptote at the input probability level. Individual curves for all seven *S*s reached the same asymptote for a given input level; there were no differences between age levels with respect to asymptotes reached. Thus the children did not tend to maximize probabilities, i.e., the curves did not converge to a probability of 1.00 but to the input levels. Thus, the children operated exactly like adults in similar studies (6, 8), in that their response probabilities approach the stimulus input probabilities. It should be noted that the response curve for an input of .60 also stabilizes at that level, indicating that children successfully discriminate .60 from chance.

Figure 2 repeats some of the information in Figure 1 in that it shows that the median response probability curves for the input level of .60 did stabilize at .60. In addition, Figure 2 shows how the distribution of response probabilities changes as a function of practice. The variance systematically decreases with practice—another indication that children can discriminate .60 from chance probability in a two-choice situation. Even the three- and four-year-olds discriminated this small deviation. Thus the threshold for probability discrimination in children would appear to be approximately the same as for adults (5).

In the data collected in Exp. I, only the slightest hint was found that the rate of learning was positively related to chronological age, and even this slight indication of slower learning for the younger children may be accounted for by their greater distractibility and inattention.

### 2. *Experiment II*

The results of Exp. II are summarized in Figure 3. Although it was felt that *S*s might divide into two or more discrete groups with distinct initial biases toward saying "happy" or "sad" and that there might be marked differences in rates of learning, the extreme effects illustrated in Figure 3 were not anticipated. One group of children showed an almost complete bias toward saying "happy" initially and actually had to *extinguish* their "happy" response *down to* the input level of .75. A second group exhibited a slight bias toward perceiving the stick-men as "sad" initially and *learned* (in the usual sense) to say "happy" *up to* the .75 input level.

### 3. *Experiment III*

Experiment III was designed as an attempt to obtain experimentally the maximizing of probability that was not found in Exp. I. It might be argued that people maximize the probability of reinforcement only if the reward is desirable enough and that merely telling *S*s "right" in Exp. I was not sufficient incentive. Since *E*'s verbal approval in Exp. I might not have been sufficient incentive for the children to maximize the probability of reward, the effect of increasing the incentive was studied. For this purpose a candy reward was also given for being correct.

The results of Exp. III are summarized in Figure 4. The increased incentive did produce a reward maximizing tendency, but it was only partially effective. Over all *S*s, the asymptote for the .75 input probabilities used in this experiment did shift upwards significantly, but the most interesting effect was an interaction between age levels and change in incentive. The output curves for the seven- and eight-year-olds converged to 1.00 (perfect maximization of reward), those for the five-year-olds stabilized at approximately .90 (a compromise between input level and maximization of reward level), and those for the three- and four-year-olds stabilized at the input level again. This interaction and general upward shift will be discussed in the section to follow.

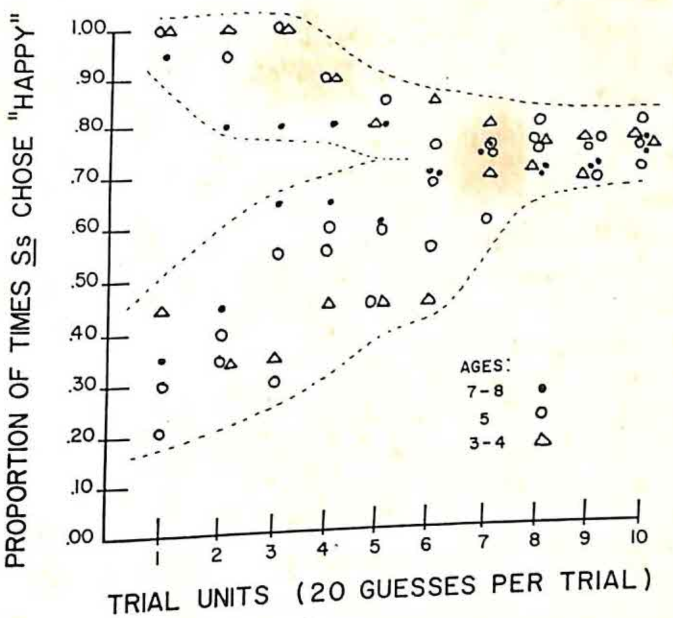

FIGURE 3

Probability Learning Curve for the Characteristic "Happy" at an Input Level of .75

## D. Discussion

In Exp. I it was found that children operate in a manner similar to adults in that (*a*) their response probability curves reach an asymptote at the stimulus input level; and (*b*) their discrimination thresholds for probabilities deviating from chance are approximately the same as those of adults. In the present experiments, these two variables did not vary as a function of age. This experiment also indicated that children can and do learn relative probabilities of stimulus events occurring in their environments.

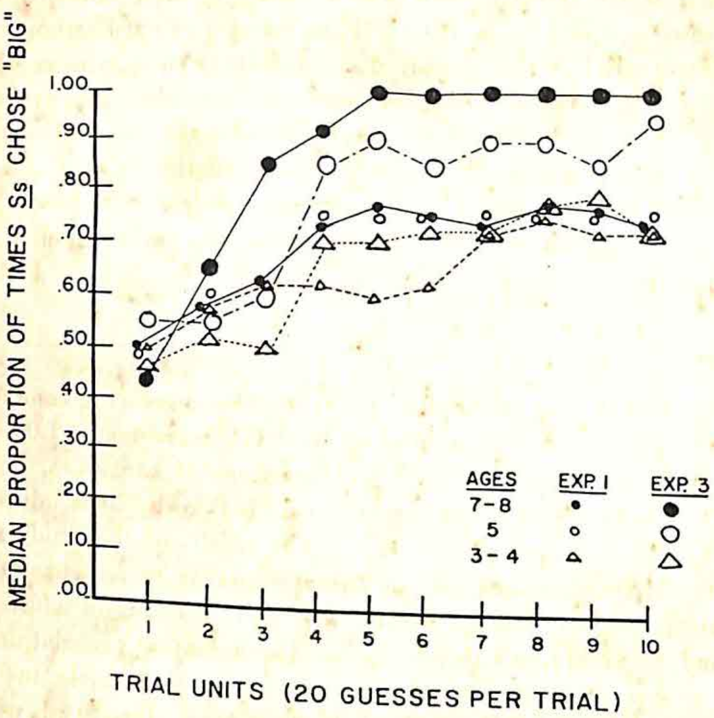

FIGURE 4
Median Probability Learning Curves at an Input Level of .75 Under Conditions of Verbal Approval (Exp. I) and Verbal Approval With Candy Reward (Exp. III)

In Exp. III it was observed that there is a marked tendency to maximize probability of reward when the value of the reinforcement (being told "right") is increased (to being told "right" and being given a piece of candy). This result could be predicted from several theoretical points of view. For example, in the Hullian model (7) an increase of drive multiplied by a specific level of habit strength (D. $_{S}H_{R}$) should raise the level of performance.

The results of Exp. III also indicated a relationship between age level and tendency to maximize probability. There are at least two possible explanations of this result, which can perhaps be differentiated in future research. It is possible, first, that the results reflect a "learning to learn" phenomenon in that children learn how to maximize the probability of reward with increased age and maturation. Thus when a reward strong enough to evoke this potential is introduced, the older children maximize the probability, the younger children do not perfectly maximize the probability of reward, and the youngest children cannot maximize the probability of reward. An alternative and perhaps complementary explanation is that the incentive value of candy increases with chronological age, at least for those age groups studied. Thus the same reward stimulus cannot be assumed to elicit the same drive at different age levels.

## E. SUMMARY

Three experiments of probability learning in children were conducted. In the first experiment, four input probability levels were used—1.00, .90, .75, and .60—with the stimuli consisting of drawings of kangaroos, horses, and stick-men. The task was to guess whether a "big" or a "little" drawing was coming up next in a deck of 20 cards. The results of this study suggested that (*a*) children's response output curves converge to an asymptote equal to the input probability level, a parallel of the performance of adults in other studies; and (*b*) children's thresholds for discriminating probabilities deviating from chance in a two-choice situation appears to be close to those obtained from adults. In Exp. II stick-men figures were used which had "happy" or "sad" faces, with an input probability of .75 for "happy" and .25 for "sad." The children split into two discrete groups—one showing an extreme initial bias toward perceiving "happy" stick-men and extinguishing down to the .75 input level and one showing a slight initial bias toward "sad" and learning up to the input level. In Exp. III a candy reward was added to verbal approval for correct guesses of "big" and "little" stick-men (input probability .75 for "big" and .25 for "little"). The seven- and eight-year-olds maximized probability, i.e., their response probability curves converged to 1.00; the five-year-olds stabilized at .90; and the three- and four-year-olds converged to the .75 input level for "big."

## References

1. Brunswik, E. Probability as a determiner of rat behavior. *J. Exp. Psychol.*, 1939, **25**, 175-197.
2. ———. Organismic achievement and environmental probability. *Psychol. Rev.*, 1943, **50**, 255-272.
3. Brunswik, E., & Herma, H. Probability learning of perceptual cues in the establishment of a weight illusion. *J. Exp. Psychol.*, 1951, **41**, 281-290.
4. Bush, R. R. Mathematical models for two-choice experiments. Talk delivered to the psychological colloquium, Univ. Illinois, 1954.
5. Detambel, M. H. A test of a model for multiple-choice behavior. *J. Exp. Psychol.*, 1955, **49**, 97-104.
6. Hake, H. W., & Hyman, R. Perception of the statistical structure of a random series of binary symbols. *J. Exp. Psychol.*, 1953, **45**, 64-74.
7. Hull, C. L. Principles of Behavior. New York: Appleton-Century, 1943.
8. Humphreys, L. G. Acquisitions and extinction of verbal expectations in a situation analogous to conditioning. *J. Exp. Psychol.*, 1939, **25**, 294-301.
9. Jarvik, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. Exp. Psychol.*, 1951, **41**, 291-297.

*Educational Testing Service*
*Princeton, New Jersey*

*The Journal of Genetic Psychology*, 1957, 90, 33-57.

# A STUDY OF SPATIAL LEARNING IN A LINEAR MAZE*

*Department of Psychology, University of Texas*

Frederick R. Fosmire and W. Lynn Brown[1]

## A. Problem and Literature

This investigation of spatial learning in the white rat was conducted in an attempt to answer the following questions:

1. If the rat is not "forced" to respond by turning right or left at a choice point, will it, in fact, behave at the choice point in a manner which indicates it "recognizes" where it is in the experimental room and/or maze?

2. Of three possible sources of orientational stimuli (extra-maze, intra-maze, and kinesthetic), which are necessary, alone or in combinations, for continued correct performance in the linear maze after the maze has been learned under constant extra-maze, intra-maze, and kinesthetic conditions?

3. Of three possible sources of orientational stimuli (extra-maze, intra-maze, and kinesthetic) which are necessary, alone or in combinations, for learning of the linear maze?

4. How do the different sources of orientational stimuli, alone and in combinations, contribute to the continued elimination of errors in the linear maze by the white rat after the maze has been learned under constant extra-maze, intra-maze, and kinesthetic conditions?

5. How do the different sources of orientational stimuli, alone and in combinations, facilitate reduction of errors in the linear maze by the white rat?

Since Small (14) first studied the learning process by observing the maze performance of rats, psychologists have been interested in ascertaining the cues which contribute to learning the maze habit and which facilitate maze running once the pattern is learned. Small set the pattern when he concluded that "neither sight nor smell, but probably touch and kinesthesis, furnished the essential data for the rat's maze running." In 1907 Watson (18) stated flatly that "vision played no part in the maze association." He also ruled out cues from olfaction, audition, gustation, and cutaneous sensations. He states: "Are we not, then, forced to conclude that the white rat

---

*Received in the Editorial Office on August 6, 1955.

[1] The cost of this investigation was met in part by a grant to the senior author by the Research Institute of The University of Texas.

makes the correct turns in the maze on the basis of the intraorganic sensations—the kinaesthetic sensations coupled with the organic probably, and possibly with the static."

The first psychologist to openly disagree with the positions taken by Small and Watson was Hunter (10). He argued:

> . . . If the conventional account of the reactions in a spatial maze were correct, the rat should be able to learn a temporal maze of almost any complexity. It is said that in the former maze, the kinaesthetic cue from running one segment controls the runinng of the succeeding segment, so that a chain of proprio-ceptive activity results. The work on the temporal maze indicates that it is all but impossible to set up a mere temporal sequence of kinaesthetic processes with the rat. Running the spatial maze therefore, must require cues which have space location as well as temporal position. In other words the rat must recognize in terms of space where he is in the maze. These spatially located cues may be anything but a kinaesthetic complex. This is not to say that a characteristic kinaesthesis will not arise in certain portions of the maze. It is to point out that the space location of the kinaesthesis is due to non-kinaesthetic processes (contact, e.g.). The present tests have indicated that the rat has practically no capacity to set up habits where the sensory complexes succeed each other merely in time. It is possible that a rat might learn a space maze requiring simple alternation and then run it in terms of kinaesthesis. The animal however easily masters the ordinary maze where the choices may be in any combination. How can it use the kinaesthesis connected with a left turn at one time to initiate a turn to the right and at another moment to initiate a turn to the left? Our tests have indicated that this cannot be done unless spatially arranged cues are available (pp. 16-17).

It was apparent to perceptive psychologists in 1920 that results obtained from running rats in alley mazes could not be compared directly with results obtained from animal performances on elevated mazes. But for 30 years more evidence emphasizing the importance of cues from one or more of the modalities was accumulated through experimentation and reported in the journals. Fewer psychologists took the position of Watson in attributing the all-important rôle in maze learning to kinesthesis. Weaver and Stone (19) found that, in an alley maze, there was no significant difference between blind and normal rats. They report that "blind animals make approximately the same errors with the same relative frequency as the normal controls."

In 1929 Hunter (11) resumed his attack on the proprioceptive theory of maze learning. He ran his animals on double-alternation bi-dimensional spatial maze and on a tri-dimensional double-alternation spatial maze. Five rats which had learned the tri-dimensional double-alternation spatial maze were run on the bi-dimensional double-alternation spatial maze. There were recorded only 13 correct runs out of more then 550. Hunter interpreted these results as pointing up the limitation of kinesthetic cues in maze learning. He admitted that proprio-ception may play a significant rôle in the learning of some mazes but not in learning spatial mazes.

Dennis concurred with Hunter in the same year. Using a walled maze, of which the walls could be removed, Dennis (3) reported that:

> . . . the results of the present experiments indicate that the rat cannot make responses of the degree of accuracy of spatial adjustment required by the maze on the basis of proprioceptive processes. . . . To a large extent and perhaps entirely, the seeing rat can dispense with these cutaneous stimuli and utilize vision to supplement the proprioceptive processes (p. 86).

By requiring rats to swim through the maze Macfarlane (12) changed the pattern of kinesthesis which may have been established during the learning trials by running. Macfarlane observed that after the first few swimming trials there was no significant difference between running performance and swimming performance. He concluded: ". . . We may speculate concerning the part played by kinesthesis in determining how the food box shall be reached, but we may safely deny it the power to tell the rat *where* he shall go to reach it" (p. 304).

Walton (17) used a sectional shifting serial choice reaction apparatus, each unit of the Yerkes type, in studying the effects of vision and kinesthesis in maze learning. The results of his second experiment ". . . cast grave doubt on the validity of the so-called purely kinaesthetic habit of maze running." Walton further concludes: "Visual habits can be so firmly established as not to be abandoned for the hypothetical internal muscular cues even when ample opportunity for such transfer is offered. . . . It is safe to conclude that the rat has many resources at his command for learning and running a maze; that one which is found most serviceable under any set of conditions is the one which becomes established and, being so established, is not easily upset" (p. 74).

The problem of sensory control in maze learning was, to some extent, clarified by Honzik (9) in a systematic study of the problem. He was decisive in relegating the dominant rôle in spatial maze learning to vision. But the problem still remained: What aspect of vision is so important? As Honzik put the question:

> . . . In view of the large part played by vision, the question arises, are all visual stimuli equally important, or are there groups of stimuli that assume an exclusive or almost exclusive rôle? Stating the question in concrete terms: Were the rats guided by the general visual environment surrounding the maze, by the gestalt, as we may say, of the room; or were they rather guided by more specific visual stimulation, as for example, by seeing the ends of blinds from the choice points? (p. 19).

In his attempt to answer this question Honzik used 90 rats, 48 experimentals and 42 controls. He screened the blind alleys and the true pathway with black cloth curtains. When he compared the performance of the two groups he found that the group which had run on the curtained maze was distinctly retarded. This he partially explained in terms of the limitation of visual cues within the maze. As the author pointed out:

> . . . Although the curtains stood only six and one-half inches above the pathways, 28 such curtains must have diminished in some degree the view of the room and general surroundings. We are forced to conclude, therefore, that the observed retardation could have been due in part at least to this fact. But it will be seen in the next experiment that the total exclusion of visual stimuli from the room . . . caused comparatively little retardation. From this it may be inferred that the partial exclusion of *room stimuli* by the curtains could have had only a small effect (p. 21).

Honzik reached the tentative conclusion that extra-maze stimuli are not as important as intra-maze stimuli in spatial-maze learning. He realized the limitations of his experimental setup and was justly conservative in his remarks.

> . . . In general it may be said that our efforts to separate visual stimuli into two distinct groups were not entirely successful and therefore the relative importance of the two groups was not clearly determined. It is possible, however, to draw the conclusion with some certainty that both intra- and extra-maze visual stimuli are used in learning. Further investigation in this field will undoubtedly yield interesting facts on the use of visual stimuli and thereby shed light on the nature of learning (p. 23).

The experimentation reported by Honzik established the main problem of the specific rôle of vision on spatial maze learning. Experimentation by other psychologists continued to minimize the importance of some of the other modal cues. Evans (4) found that kinesthesis contributes no essential cues to establishment of the maze habit.

By the employment of a tunnel maze, Brown (1) found supporting evidence for the theory that in enclosed mazes the dominant guiding cues are maze contained. When the maze was shifted 5′ there was no significant increase in error scores. Even when the maze was moved into a room which provided a homogeneous extra-maze visual environment and run under conditions of bright illumination the experimental rats made error scores which were insignificantly different from previous error scores made in the experimental room.

Tolman, Ritchie, and Kalish (15) began a series of studies in spatial learning in 1946. Animals were trained on a pathway without culs which led from a table top to a goal box where they were rewarded. A 5-watt light behind and above the level of the goal box was the only illumination in the experimental room. On the test trial the entrance to the training pathway (Pathway *CD*) was blocked and 18 new pathways led from the table top, nine from each side of Pathway *CD*. Path No. 6 presented the most direct pathway to the former goal location. Of the rats which made a choice before six minutes lapsed, 36 per cent selected Pathway 6. It was pointed out by Tolman, Ritchie, and Kalish that this could not be explained in terms of generalizations since the visual cues from the light source are practically the same for Pathways 5 and 7 as they are for Pathway 6, and only 10 per cent of the rats chose either Pathway 5 or 7. They also explain that it is not a matter of conditioning to run toward a light source.

> . . . In answer to this criticism we should first explain that we are not asserting that rats can exhibit such orientation behavior when there are no cues or landmarks present. We believe that such choices can only be made when there are distinctive stimuli in the environment which enable the rat to judge its own location relative to other places in the environment. The light, we believe, performed such a function and was not a mere conditioned stimulus, as such a criticism would suppose (p. 20).

In their second experiment on spatial learning Tolman, Ritchie and Kalish (16) attempt an explanation of what a rat "learns" in learning a simple *T*-maze. They say:

> . . . 1. Such training may have produced a disposition in the rats to run on a path which has certain specific characteristics (e.g., knotholes of such and such a pattern, or the like) and to avoid running on all paths which have certain other specific characteristics. 2. Such training may have produced a disposition to turn right whenever they come to the choice point. 3. Finally, such training may have produced a disposition to orient towards the place where the food is located (e.g., under the window, to the left of the radiator, etc.) (p. 221).

It is the authors' contention that there is at least one other possibility. If rats are capable of "cognitive maps" then it is possible that a rat could be space oriented *within* the maze as well as within a room providing extra-maze visual cues. (This could be tested by varying the extra-maze visual cues in relation to the maze location such that the rat could not perform correctly with a "cognitive map" based on extra-maze visual cues.)

Gentry, Brown, and Kaplan (6) found cause to disagree with the reported experimental results and interpretations of Tolman, Ritchie, and Kalish in their first experiments. Gentry, Brown, and Kaplan used apparatus and procedure which essentially duplicated that of Tolman, Ritchie, and Kalish. The results obtained by Gentry, Brown, and Kaplan were not similar to those obtained by Tolman, Ritchie, and Kalish, and Gentry, Brown, and Kaplan remark:

> . . . It is evident from the results presented in this study that different positions of the light during preliminary training and on the test trial did not influence the performance of the animals on the test trial (p. 319).

They conclude:

> . . . While we found no evidence, even in the case of the group given nine days of training—five days more training than Tolman, Ritchie, and Kalish gave their rats—we would argue that even if their rats did learn the locus of the food box, such learning is in itself inadequate to explain their 100 per cent correct choices. In view of the close proximity of Paths 5 and 7 to Path 6 and the absence of distinguishable features at the choice points in the apparatus, it seems very dubious to us that the 100 per cent correct choice of Path 6 is explicable simply in terms of the assumption that the correct choice rats had learned the location of the food box. Under such conditions the best that one could expect would be a clustering of choices around Path 6 (p. 321). . . . Consequently, since our results for both animals with training and without training are in part explainable in terms of an artifact of the experimental apparatus (the presence of Alley *CD*), we wonder if their results are not explainable in terms of some artifact of their experimental set-up, and hence if there is any need for them to appeal to the senior author's expectancy theory in its inferential form (pp. 321-322).

A second test on spatial learning was conducted by Gentry, Brown, and Lee (7). Eighty white rats were given one run a day on the maze apparatus shown in Figure 1. After 11 days of training the apparatus depicted in Figure 2 was introduced. Each rat was allowed to run from *A* to *B*, explore the table top (which was frequently rotated), enter the blocked pathway, and choose a final pathway. Approximately 31 per cent of the animals chose Pathway 2; 21.25 per cent chose Pathway 3, and only 5 per cent chose Pathway 8 (which led directly toward the light at *H* and in the direction of placement of the food box during the first 11 days of training).

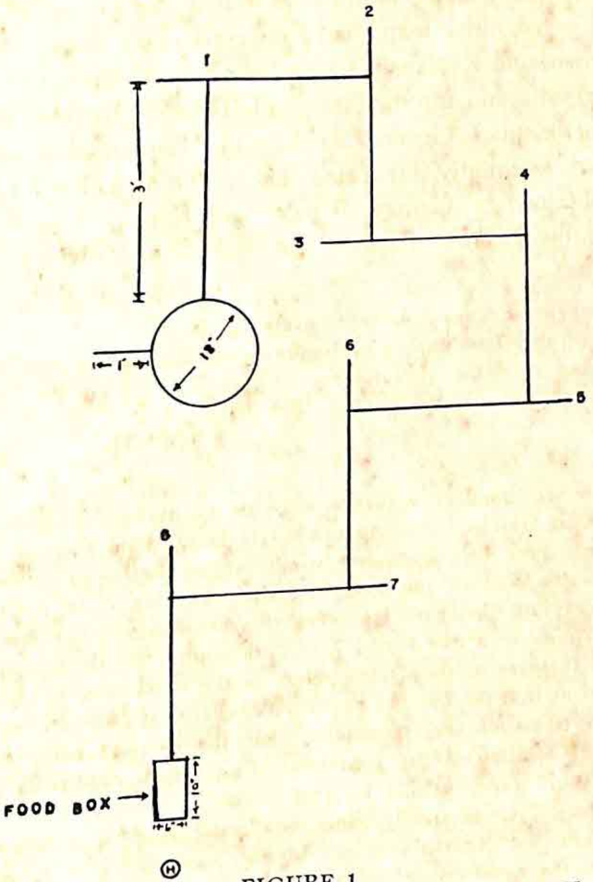

FIGURE 1

MAZE USED BY GENTRY, BROWN, AND LEE IN PRELIMINARY TRAINING

Gentry, Brown, and Lee say:

> . . . The behavior of the animals on the test trial gave us no reason for believing that the light (or locus of the goal box) had any material bearing on the performance of the animals or reasons for thinking that they had acquired even the rudiments of a general directional orientation with reference to it (or them). . . . The animals probably

> acquired a directional orientation and tendency to run in a given direction from a given part of the table top or learned that the appropriate point of departure from the table top was a given area and that progress in a given direction from this area of departure was a step in the action sequence leading to the goal box (p. 317).

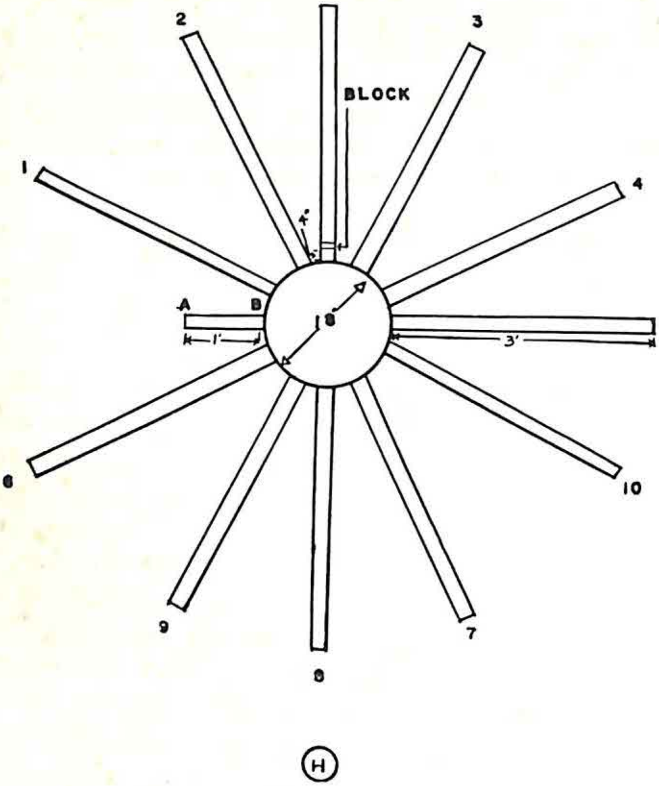

FIGURE 2
APPARATUS USED BY GENTRY, BROWN, AND LEE ON TEST TRIALS

This last supposition is certainly not negated by the Ritchie (13) report that over half of his experimental animals exhibited direction dispositions but that less than 10 per cent acquired a place disposition.

The problem of how rats learn spatial maze problems is solved somewhat (and complicated somewhat) as a result of experiments conducted by Brown and Humphrey (2). They tested the hypothesis that place learning is the learning of a definite concrete situation in space. The Control Group ($N = 17$), during seven days of training, was given one run a day on each of the positions shown in Figure 3 (a total of eight runs a day).

(Only one *T*-maze was assembled during training.) All animals ran without error on the seventh training day.

A series of tests was then given to determine the factors involved in learning the problem. In the first test the maze was rotated through 45 degrees, and the animals ran Positions 1-8 with 97 per cent correct choices. When compared with the initial day of training a *t* of 3.14 was found. In testing for place disposition and response disposition, the animals were run in Positions 1-8 in a randomized sequence. The percentage of correct runs on this test was 93. For a third test the controls were shifted to a room with homogeneous extra-maze cues where they ran with 91 per cent correct choices.

Brown and Humphrey utilized three experimental groups in addition to the group just described. Experimental Group 1 ($N = 17$) was given training on Positions 1 and 2 only for seven days. On the eighth day these animals were tested on all eight maze positions and made 74 per cent correct runs on Positions 3 through 8. Experimental Group 2 ($N = 17$) was given seven days of training on Positions 1-8 in a homogeneous dome room. Following training they were tested under drastically reduced illumination and were able to run with only 54 per cent correct runs. Experimental Group 3 was trained with reward at the four wall positions rather than the center of the room. When the maze was rotated 45 degrees, with the four reward positions now in four new positions, the animals ran Positions 1-8 with 96 per cent correct runs. Brown and Humphrey: "We can, therefore, say that a definite concrete situation is not required in spatial learning." When they comment on the few errors made in a novel environment which is dissimilar to the training environment they write:

> . . . Inasmuch as the only common feature of the two environments was the relation of the choice points to the walls, it would seem that what had appeared to be place learning of a finite location in space was in reality a learning of the food position in relation to a non-specific wall surface.

Feddersen and Brown (5) trained rats for 10 days on starting Positions 7, 8, 3, and 4 of the Brown and Humphrey maze (see Figure 3), and then ran them for 10 more days on starting Positions 1, 2, 5, and 6. The 18 animals made 32 errors on the first day of training and four on the first day of the second 10-day period. This difference is significant beyond the 1 per cent level of confidence. The 90 degree rotation of the maze between training periods did not affect maze performance. The place of the goal box remained constant throughout both training periods.

## B. MATERIALS AND METHODS

### 1. *Maze and Room*

The linear maze and experimental room used throughout this experiment are shown schematically in Figure 4. It was constructed of dressed pine and was painted glossy black. The maze stood 15″ from the floor, had running tracks 2″ wide, and measured 24″ between all choice points, 24″ from the start to the first choice point, 24″ from the last choice point to the end of the maze, and the cul-de-sacs were 24″ long. Dummy food boxes were placed at the terminus of all arms except the arm which was to be correct for any particular trial. The dummy food boxes were painted black, measured 10 x 6 x 5 inches, and resembled the genuine food boxes except for the missing floor. In some of the experiments the starting path was lengthened by adding elevated units 24″ long, 15″ high, and with a 2″ running track. These were also painted flat black.

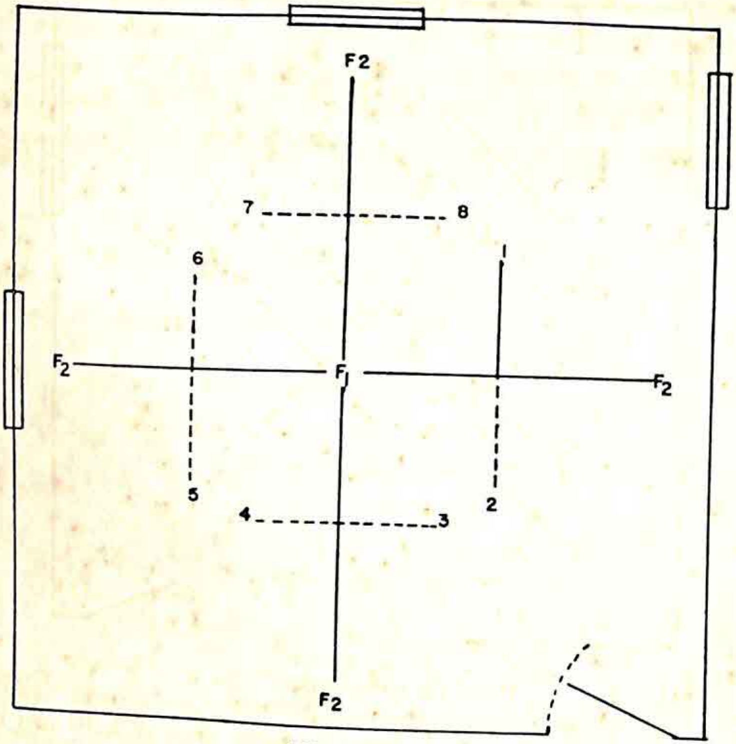

FIGURE 3
FLOOR PLAN OF MAZE USED BY BROWN AND HUMPHREY

The arms jutting to the right of the main pathway were designated as even-numbered culs and the arms leading to the left of the main path were designated as odd-numbered culs. The 24″ section beyond the last choice point was called the "end unit." The 24″ section before the first choice point was called the "starting section."

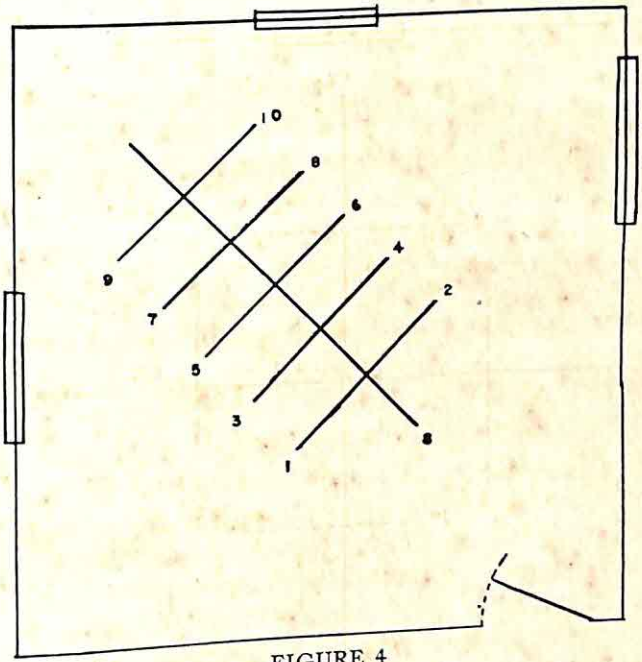

FIGURE 4
Maze in Experimental Room

The maze was assembled in a laboratory room 17 x 17 x 10 feet in dimension. The different maze settings, which will be described later, were along an imaginary axis from the Northwest to the Southeast corners of the room. A reference point along this axis, where the Northwest-Southeast diagonal intersects the Northeast-Southwest diagonal, is referred to as the point "*X*." This point was carefully marked in the room but was not available to the rats as a cue.

## 2. *Preliminary Training and Scoring*

None of the animals was given preliminary training. All animals were placed on a feeding schedule for at least six days prior to the first trial run. At the time they were fed they were also handled so that when they were introduced to the maze, though they had had no previous maze experience, they were tame and fairly active.

Body length entry into any of the arms constituted an error. If an animal entered the end unit an "*E*" error was scored. A retrace was scored if an animal returned in the direction of the starting section before it reached the correct choice point or if it continued past the correct choice point returning from errors in the latter part of the maze. No animal was allowed to retrace to an earlier choice point, or more than 24″ if the retrace occurred in the starting section.

### 3. *Animals*

The animals were experimentally naïve and were from 75 to 180 days old at the beginning of the experiment. They were from the Comparative Psychology Laboratory colony.

### 4. *Method*

The correction method was utilized in the experiment.

## C. PROCEDURE

The animals were housed outside the experimental room and were brought into the experimental room individually each time they were given a maze trial. Four runs daily were given all subjects.

While the experiment was in progress, the experimental room contained nothing but the maze, food boxes, food box supports, the animal running the maze at the time, and the experimenter. Each animal was allowed to eat about 60 grains of wet mash in the food box, after which it was returned to its home cage.

### 1. *Experiment 1B*

Twenty rats, 10 males and 10 females in Group C1, were given four runs daily with the maze in a constant position (Choice-point 2 over *X*) and with a genuine food box at the end of Arm 4. As they reached the criterion of at least two consecutive errorless days (8 perfect runs) they were divided into matched subgroups, Groups IIIB and IVB. For the rats in Group IIIB the correct choice-point was in the *same place in the room* every day. Extra starting units were added when necessary to maintain the distance of 10 feet from start to correct choice-point as the maze was shifted along a Northwest-Southeast axis. Thus the extra-maze visual cues and the kinesthetic cues were constant but the correct intra-maze choice-point was always changing. The ratio of total errors to total runs fluctuated between 2.1 and 0.3 for nine days after which the error ratio dropped from 2 on the ninth day to 0.2 on the 15th day.

For the rats in Group IVB the maze was shifted in exactly the same manner as for Group IIIB, i.e., different choice-points were over the $X$ on each run, but with Group IVB rats the *same place in the maze* was always correct. Distance run was also kept constant at 10 feet so that kinesthetic cues were available if the rat could utilize them.

### 2. *Experiment 2B*

Twenty rats, 8 males and 12 females, in Group C3, were given four runs daily with the maze in a constant position (Choice-point 2 over $X$) and a genuine food box at the terminus of Arm 4. As the rats reached criterion of two consecutive errorless days, they were divided into matched Subgroups IIIB and VIB. With the maze in a constant position in the room and with Arm 4 always the correct pathway to food, units were added as necessary to modify distance run in a systematic manner. Thus, for animals in Group IIB, extra-maze visual cues and intra-maze cues were constant but possible kinesthetic cues were variable, and, henceforth, unreliable for use in solving the maze problem.

Group VIB animals were given trials under conditions of shifting extra-maze cues and changing distances from start to correct choice-point. Only intra-maze cues were invariable.

### 3. *Experiment 3B*

Eighteen rats, 10 males and 8 females, constituted Group C2. They were run four times daily on the maze with extra-maze cues, intra-maze cues, and kinesthetic cues constant. As the rats reached criterion of two consecutive errorless days (at least eight consecutive errorless runs) they were divided into matched subgroups, seven rats in Group VB and eight rats in Group VIIB.

Rats in Group VB were required to run the maze under conditions in which only extra-maze conditions were constant. The maze was shifted along its axis and units were added as necessary to vary systematically intra-maze cues and kinesthetic cues from distance.

For rats in Group VIIB the maze was shifted and units added in such a manner as to maintain unchanging intra-maze cues and cues from distance from start to correct choice-point. The animals could not rely on extra-maze visual cues as the correct choice-point changed in locus within the room from one run to the next.

### 4. *Experiment 1A*

In Experiments 1B, 2B, and 3B the experimental animals first learned the linear maze with all available extra-maze visual cues, intra-maze cues, and kinesthetic cues of distance from start to correct choice-point constant. As the rats reached the criterion of two consecutive errorless days they were divided into six different subgroups which were given trials under conditions of reduced cues.

In Experiments 1A, 2A, and 3A the experimental rats were given trials, from the beginning run until the final run, under conditions of limited available cues. Observations of the performance of Groups IIA and VIIA comprised Experiment 1A.

The rats in Group IIA were given trials in an experimental situation in which distance from start to correct choice-point was varied systematically. Extra-maze and intra-maze relations were kept constant.

Group VIIA was run under conditions in which only kinesthetic cues of distance run were unchanging. Distance from start to correct choice-point was kept constant at 10 feet.

### 5. *Experiment 2A*

Experimental conditions for Groups IIIA and IVA were similar to those in Experiment 1A in that distance from start to correct choice-point was kept constant at 10 feet. For Group IIIA the maze was shifted along its axis between trials for each animal but the correct choice-point was always over $X$. Extra-maze cues were changing and, therefore, unreliable for rats in Group IVA; only intra-maze and kinesthetic cues were constant.

### 6. *Experiment 3A*

The correct choice-point was always over $X$ for Group *VA* but intra-maze and kinesthetic cues were unreliable. The correct choice-point was constant within the maze for rats in Group VIA but it was always changing within the room. Distances from start to correct choice-point varied also for Group VIA.

## D. RESULTS

### 1. *Experiments 1B, 2B, and 3B*

When the white rat is not forced to make a differential response at a choice-point but is allowed to continue forward beyond the place in the maze and room where a right or left turn is goal-leading, it can still learn a spatial maze.

In Experiment 1B, 20 rats in Group C1 were given four runs daily with the maze in a constant position. The second choice-point from the starting arm was always located over a small "*X*" penciled on the floor representing the center of the experimental room. Food was placed in the food box at the end of Arm 4 and dummy food boxes (boxes without bottoms) were located at the ends of all other arms.

Eighteen of the 20 rats in Group C1, running under conditions which held constant extra-maze, intra-maze, and kinesthetic cues, learned the linear maze by the 40th trial. Five rats in the group reached criterion (two consecutive errorless days) at the end of the sixth day.

As the rats in Group C1 reached the criterion they were divided into matched subgroups, Group IIIB and Group IVB.

The rats in Group IIIB found that the correct choice-point was constant each day with respect to room environment (same place in room) but that food might be located at the end of Arms 2, 4, 6, 8, or 10. Distance from start to correct choice-point was held constant at 10 feet by adding two-foot maze starting units when necessary.

Group IVB animals (running to same place in maze) made less errors during the last nine days of the experiment than rats in Group IIIB, but the difference was not significantly different from chance. Apparently, the white rat can continue to run the linear maze under conditions of reduced cues as well using extra-maze space cues and cues from distance as using intra-maze cues and cues from distance.

Group C3 consisted of 20 rats given the same preliminary training as those in Group C1. As the animals reached the criterion of two consecutive errorless days they were transferred into matched subgroups, IIB and VIB.

The rats in Group IIB no longer ran the same distance each trial before reaching the correct choice-point. Extra-maze and intra-maze cues remained constant but possible kinesthetic cues were variable, and, therefore, unreliable for use in solving the problem. Maze units were added as necessary to modify distance run in a systematic manner. The ratio of total errors to total runs varied between 1.9 on the first day to 0.4 on the 15th day.

For the rats in Group VIB the maze was shifted after each run in such a way that the choice-point associated with any arm on the right side of the maze could be over "*X*." Units were added to the starting arm as necessary to insure variable distance from start to correct choice-point. The goal-box was always located at the terminus of Arm 4. Thus, the procedure was arranged that when Choice-point 5 was over "*X*," Choice-point 2 was correct and distance from start to correct choice-point might be four feet (or six feet if one additional unit was added); when Choice-point 4 was over "X," Choice-point 2 was correct but distance from start to correct choice-point might be four, six, or eight feet. By varying the number of two-foot starting units added to the regular starting arm, it was possible to make sure the rats were not learning five problems simultaneously, each based on constant intra-maze and distance signs. The ratio of total error to total runs was between 1.2 on the first day to 0.3 on the 15th day.

When a comparison is drawn between Group C3 and Groups IIB and VIB, it is found that Groups IIB and VIB made significantly more errors than Group C3. Groups C3 and IIB differed by a *t* of 4.9 and C3 and VIB by 49.7.

Experiment 3B was conducted with 18 animals in Group C2 which was later divided into Subgroups VB and VIIB. Group C2 rats were given four trials daily under the same conditions as Groups C1 and C3. Results of "*t*" tests of significant difference between the three groups (C1, C2, and C3) show that none of the differences is significant.

Rats in Group C2, as they reached the criterion of learning, were divided into Groups VB and VIIB. Group VB animals were required to extinguish any goal-leading responses based on intra-maze or distance cues if they were to travel from the starting arm to food-box in the shortest possible time. The choice-point "*X*" was always correct, i.e., a right turn at the choice-point over "*X*" always led to food. Any choice-point could be over "X" and distance was variable from two feet to 10 feet.

The animals in Group VIIB could make errorless runs only by extinguishing goal-leading responses based on intra-maze and extra-maze cues and by relying solely on constant cues associated with distance run on each trial. Distance from start to food-box was maintained at 10 feet; Choice-points 1 through 5 could be over "*X*"; and, depending upon which choice-point was over "*X*," any even-numbered arm could lead to food. For example, when Choice-point 3 is over "*X*," Arms 4, 6, 8, or 10 could lead to food, depending upon the number of units added to the regular starting arm.

The performance of Groups VB and VIIB differed significantly from that of Group C2. Groups C2 and VB differed by a *t* of 10.1 and C2 and IVB by a *t* of 9.9. Group VB made significantly more errors on Days 7 through

15 than Group IIIB, but the differences in error scores between Group VB and Groups IIB, IVB, VIB, and VIIB might be expected on the basis of chance alone. Group VIIB made significantly more errors than Groups IIB, IIIB, IVB, and VIB. The experimenters were purposely cautious in computing the "*t*" score between Groups VB and VIIB and did not consider them matched groups for statistical purposes. This, of course, yielded a smaller "*t*" than if the correction for correlation between groups had been made, but it was thought that the more conservative estimate of difference was in order.

A general statement relating the nature of errors making up the total error score for each day should be made. After the second day (eight trials) almost all of the errors made by animals in Groups C1, C2, and C3 were right-turn responses, almost all of these being anticipatory. Almost all of the errors observed for Groups IIB, IIIB, IVB, VB, VIB, and VIIB were right-turn responses.

### 2. *Experiments 1A, 2A, and 3A*

The procedure for testing the white rat's ability to localize the food box in a linear maze, without all extra-maze, intra-maze, and distance relations unchanging, if the animal is not given the opportunity to learn the maze under constant extra-maze, intra-maze, and distance conditions before being tested in situations yielding fewer environmental "signs," but given four trials daily under conditions of variable extra-maze, intra-maze, and distance cues, or combinations of these, from the beginning, are very much like those utilized in Experiments 1B, 2B, and 3B. Whereas, in Experiment 1B, the rats were required to make at least eight consecutive errorless runs before they were tested either in a situation in which cues from room "place" and distance from start to correct choice-point were as before and unvarying but in which maze "place" changed in a systematic manner, or in a situation in which maze "place" and distance from start to correct choice-point were constant as in the training trials but room "place" varied according to a preconceived plan, rats in Groups IIIA and IVA, respectively, were given trials from the first day under the test conditions just described. The difference in error scores for Groups IIIA and IVA during the last nine days of the experiment is not significant at the 5 per cent level of confidence.

Animals in Groups IIA and VIA were, in like fashion, given trials under "reduced cue" conditions from the beginning of the experiment. As Groups IIIA and IVA differed from Groups IIIB and IVB, respectively, only in that Groups IIIA and IVA were not given trials on the linear maze when all available environmental cues were constant before being tested when some cues were changing from trial to trial; similarly Groups IIA and VIA were different from Groups IIB and VIB, respectively, in the same respects.

For Group IIA, distance from start to the correct choice-point was not the same from one run to the next. Group VIA rats found that the correct choice-point was always the same place in the maze but it would change with respect to distant stimuli and, furthermore, start to correct choice-point distance was variable. Since the calculated "*t*" score between Groups IIA and VIA are not significantly different from chance, it must be concluded that, when the situations just described are instituted from the first trial to the 15th, the combination of unvarying room "place" cues and maze "place" cues are of little more assistance to the rat than maze "place" cues alone.

Rats in Group VA, if they were to learn the maze, would have to rely from the first day of the experiment solely on the correct choice-point always being located in the same place in the room. Proximal cues were changing from trial to trial as was distance from start to correct choice-point. Distance cues were the only unchanging cues from Day 1 through Day 15 for Group VIIA rats. All others varied unceasingly.

The "*t*" score between Groups VA and VIIA is 1.63 and not significant. Hence, there is no reliable difference between room "place" signs and "distance" signs when they are not supported by cues deriving from other constant conditions and are the only unchanging cues throughout the experiment.

Tests of significance of difference between Groups IIA, IIIA, IVA, VA, VIA, and VIIA show that the difference in errors made during the last nine days of the experiment between Groups IIA and VIIA could be expected on the basis of chance less than once in any hundred repetitions of the experiment, the same general statement applying to the error score differences between Groups IIIA and VA, Groups IIIA and VIIA, and Groups IVA and VIIA. The error score differences between Groups IIIA and VIA and Groups VIA and VIIA are significant at the 5 per cent level of confidence.

## E. Discussion of Results

The purpose of this study was to determine whether or not the white rat can learn a definite spatial location if it is not forced to respond differentially at the choice-point which is associated with the goal in an open, elevated maze, if the rat does learn such a problem whether it localizes the goal with respect to extra-maze (distant) environmental characteristics or with respect to intra-maze space relations, and which of the three variables (place in the experimental room, place in the maze, distance from start to correct choice-point) alone and in combinations are necessary and sufficient for learning the maze and continued errorless performance under conditions of reduced cues. Some of the findings were unambiguous; other results were insignificant or misleading.

It was definitely demonstrated that rats can learn a spatial maze when the conditions would allow them to continue forward. All of the reported studies in spatial learning were conducted with simple *T*-mazes or multiple-*T* mazes which required a right or left turn at the correct choice-point. It seems to the authors that a linear maze of the type employed in this experiment demands a more specific kind of space learning, for the rat is required, first, to recognize the correct choice-point and, second, to make the proper turn at that point.

From the reported observations of Groups C1, C2, and C3 it is evident that the linear maze, under unchanging maze-environment conditions, poses not a difficult task for the albino rat. If certain maze-environment relations are altered systematically from trial to trial the maze problem becomes more difficult but as long as one of the three variables mentioned above remains constant, the problem never becomes impossible. Even when place cues from within and without the maze are unreliable there occurs a reduction of errors with time as the rat adapts its behavior to be sensitive to the apparently crude kinesthetic cues from distance run.

In Experiments 1A, 2A, and 3A, it was shown that while a lower percentage of animals will learn the problem under any conditions of reduced cues the laboratory rat can master the maze if only room and maze place, room place and distance, maze place and distance, or maze place alone are kept constant. Over-all performance is best when room place and distance are unchanging, followed in order of adequacy by constant conditions of room and maze place, maze place and distance, maze place, room place, and distance.

The extra kinesthetic supports from constant start to correct choice-point distance apparently were more useful to the rat when coupled with constant room place conditions than when coupled with maze place constancy. The difference in errors made in the two experimental situations is not significant. The room place and kinesthesis combination is given attention for two reasons. First, the later literature on spatial learning implies a clear supremacy of effectiveness of spatial location over kinesthesis. Now it is tacitly assumed in most of the latter day reports that spatial localization rests almost solely on the rat's perception of extra-maze environment characteristics. But, though no one has emphasized the possibility of intra-maze space relations, it would surely be agreed that intra-maze space perceptions, if effective at all, would be more beneficial to the rat's learning than kinesthesis. The results do not support that position, but, instead, indicate unequivocally that whereas the difference between Groups IIA and VA (room and maze place versus room place) and the differences between Groups IIA and VIA (room and maze place versus maze place) could be explained on a non-chance basis less than 5 per cent of the time, the difference between Groups IIIA and VA (room place and distance versus only room place) is significant at the 1 per cent level of confidence. It should also be pointed out that cues from kinesthesis do not support maze place cues to the same extent. The difference between Groups IVA and VIA (maze place and distance versus only maze place) is not significant. Second, before the experiment was begun the authors predicted that constant maze place and distance relations would facilitate error reductions more than would constant room place and distance conditions. It was thought that since both intra-maze and start-to-correct-choice-point distance cues were proximal, one would reinforce the other when the rat inhibited the established reliance on distant stimuli (when room place relations were variable) but that when maze place relations became variable the rat would inhibit reliance on proximal stimuli, hence distance cues also, to favor distant stimuli. That this did not occur is obvious.

*Ad hoc* reasoning would permit one to say that the principle is: In situations in which some aspects of the environment are unreliable, the rat seeks both proximal and distant references. This statement is compatible with the results of this experiment but, of course, should not be considered seriously until proven by experiments *designed* to test the statement as an hypothesis. If the proposition proved to be factual it would partially explain the experimental findings of Small and Watson which are at odds with the results obtained by Hunter, Dennis, Macfarlane, Honzik, and many others.

The problems confronting the experimental animals in Experiments 1B, 2B, and 3B were considerably different from the ones just discussed. Under "*B*" conditions the rats first learned the maze task under constant conditions after which they were divided into matched subgroups to continue the problem under conditions of reduced cues. This is an entirely different situation from the general one devised for the "*A*" experiments. Animals in Group IIIA, for example, *never* experienced reliable extra-maze space cues in the experimental room. There is every reason to believe they perceived the situation in no way the same as rats in Group IIIB.

In the "*B*" experiments, cues resulting from constant maze place and distance conditions were significantly more facilitative to learning than any of the other conditions. A survey of the results of Groups VB, VIB, and VIIB will immediately demonstrate the previous experience on the maze under constant conditions does not make it easier for the rats to learn the maze when some of the previous references are variable and unreliable. Consider Groups IIB and VIB, matched subgroups of Group C3. Rats in Group IIB found a new problem in that start-to-correct-choice-point distance was no longer unchanging. The "*t*" score between Group IIB and Group C3 is 4.9, significant at the 1 per cent level of confidence. The "*t*" computed between Group VIB and Group C3 is 49.7, also significant at the 1 per cent level of confidence. The only defensible explanation of the difference must be made in terms of reduced cues since there is no evidence that the rats became insensitive to the environment and the difference is not attributable to chance. It must also be concluded that loss of cues from distance run effects less increase in error score than loss of extra-maze and distance cues.

Groups IIIB (room place and distance constant) and IVB (maze place and distance constant) were subgroups of Group C1. The changes in conditions for the "*B*" part of the experiment have been elaborated upon previously and it should be obvious that, under such conditions, if rats learn a spatial maze by relying primarily upon extra-maze visual cues with little regard for intra-maze cues, there would be no impairment of performance and one would expect all eight animals in Group IIIB to continue their errorless performance. The performance of the animals reveals that such an expectation was not confirmed by the results; that the rats were apparently disturbed, either by the absence of intra-maze cues, or more likely, by the interference of intra-maze cues. Under the conditions, Group IIIB rats, in order to reach the correct goal box by the shortest route and in the shortest possible time, would be required on some of the trials (those occurring when the correct goal box was located at Arms 6, 8, or 10) to ignore positive signs arising from a particular maze place (Arm 4 choice-point) and to continue forward until they could distinguish the locale on the basis of extra-maze and kinesthetic cues.

At this point a criticism could be justifiably raised. If one can assume that visual acuity in the white rat is sufficient to perceive all of the maze at one time, say from the junction of Arm 6 with the main pathway, and if the rat is capable of the cognitive processes necessary to perceive the maze as a whole (such processes being of no higher order than those required for the establishment of a "cognitive map" with respect to extra-maze visual cues), then it would surely be in order to argue that the rat is capable of two kinds of "cognitive maps": one based for the most part on extra-maze space characters, and the other independent of extra-maze characteristics.

If two kinds of maps could be formed there are two possibilities which would explain Group IIIB results. One, assume that when both extra-maze and intra-maze signs are invariable, the rat depends first of all upon remote visual cues (extra-maze) because as Hebb (8) points out: "When visual cues from the distant environment are available, they dominate behavior." Under ordinary conditions the perception of the distant environment is changed less by movements of the perceiver than perception of closer objects in the environment. Therefore, because of past experience, remote cues assume a primacy and, under ordinary conditions, are utilized in behavior to a greater extent than proximate cues. Assume further that when the distant environment becomes unstable, a second "cognitive map" can be established, e.g., one independent of remote space characters. If the assumptions just made correspond to actuality, one would expect a period of "re-learning" for rats in Group IIIB after the intra-maze environment becomes capricious.

There is a second possibility which, though not differing from the first in any serious aspects, might be more acceptable to some theorists. Assume, now, that under constant environment conditions, the rat establishes two "cognitive maps" of the environment corresponding to the two "maps" mentioned above. Assume further that under ordinary conditions, though either would be equally useful in learning the linear maze, a dominant reliance in past experience on the greater relative invariance of the remote environment has resulted in a utility priority for the extra-maze-anchored map. Then, when interference is introduced into the maze situation by changing intra-maze signs, the animal makes more errors until the substructure (maze anchored map) is extinguished or inhibited.

The reader apprehends that it is not necessary to explain the results obtained in Experiments 1B, 2B, and 3B; by the same principles that give meaning to the "*A*" experiments results. Different procedures can lead to the development of different kinds of solutions. In addition, the reader should also realize that in comparing the value to the rat of the three environment supports discussed there are plural ways of making an assessment. If one considers the important test of learning only the percentage of animals who reach criterion, he would likely get a different picture than if he considered average errors per animal. This is not evident in the "*A*" experiments but extremely noticeable in the "*B*" experiments. In the "*A*" experiments the conditions rank, in order of utility for learning a linear maze, as measured by the percentage of animals reaching criterion: (*a*) Room place and distance, (*b*) room place and maze place, (*c*) maze place and distance, (*d*) maze place alone. If average errors per rat are considered as the test of the importance of the three variables discussed, then the rank is as follows: (*a*) room place and distance, (*b*) room place and maze place, (*c*) maze place and distance, (*d*) maze place alone, (*e*) room place only, and (*f*) distance. The rank-order correlation between the two criteria is almost a perfect one, being reduced only in that no animals in Groups VA or VIIA made four errorless runs on two consecutive days.

In the "*B*" experiments there is less agreement between the two ways of determining importance of the variables. When the percentage of rats reaching criterion is used as the index of importance, the variables rank: (*a*) maze place and distance, (*b*) room place and maze place, (*c*) room place and distance, and (*d*) room place only.[2]

With respect to average errors for each rat the variables rank in importance: (*a*) room place and distance, (*b*) room place and maze place, (*c*) maze place only, (*d*) room place and distance, (*e*) room place, and (*f*) distance. Again, the two combinations most useful to the albino rat in learning or retaining the maze problem are unchanged in rank with either criterion.[3] If one considered only the first criterion of importance, it would appear that extra-maze space relations have greater meaning for white rats in spatial mazes than do intra-maze space relations. If one considered only the second criterion, it would appear that the opposite is true. Probably the only safe conclusion would be that, under "*B*" conditions, both distant and proximate cues are important.

---

[2] Room place can be included in this ranking only if the reader understands that the three animals which reached criterion did so on the sixteenth day, one day after the experiment should have terminated. Three rats made four errorless runs on the fifteenth day and they were given trials on Day 16 to see if they could repeat the previous day's performance. All three reached criterion on the extra day and are conditionally included in the ranking.

[3] Some might be tempted to think of the "*B*" experiments as tests of retention, but the writers would disagree. When any one of the three variables under study is changed the maze problem is essentially a new one. Though the maze and room remain the same, the relations are different or the situation is different, a matter of first importance in subhuman learning.

## F. Summary and Conclusions

These experiments demonstrate that the albino rat does not have to be "forced" to respond differentially at a choice-point in an elevated open maze but will learn to turn correctly at the bifurcation associated with reward even if the original pathway continues forward beyond the bifurcation.

Further evidence for spatial learning was obtained but it seemed to indicate that, under particular conditions, intra-maze space relations are very important to the rat in localizing the goal box. It was postulated that maybe two kinds of "cognitive maps" are established in the rat: one based upon extra-maze or distant visual cues and the other with respect to intra-maze or proximate visual cues. It was further postulated that when the albino rat's experiences in the maze-room environment, for the very beginning, occur under conditions of variable and consequently unreliable maze-room relations, the rat will use a combination of distant and proximate visual cues more quickly and more effectively for maze learning than it will use combinations of proximate cues only or distant cues only. But if the rat first learns the maze under a constant maze-room relationship and is then subjected to variations in the relationship it will not rely on extra-maze or distant visual cues any more readily than extra-maze or proximate cues.

These experiments also indicated that kinesthetic cues of distance are of considerable importance in maze learning *when* they support extra-maze or intra-maze visual cues. By themselves, kinesthetic cues aid the rat but little in finding the goal box.

The results are basically in agreement with the findings of Brown, Gentry, Hebb, Honzik, and Hunter.

## References

1. Brown, W. L. The effects of a common start-finish locus on orientation and behavior in a multiple-*T* maze. *J. Comp. Psychol.*, 1946, **39**, 331-338.
2. Brown, W. L., & Humphrey, C. E. Generalization in spatial learning. *Amer. J. Psychol.*, 1955, **68**, 396-408.
3. Dennis, W. The sensory control of the white rat in the maze habit. *J. Genet. Psychol.*, 1929, **36**, 59-90.
4. Evans, S. The rôle of kinaesthesis in the establishment and control of the maze habit. *J. Genet. Psychol.*, 1936, **48**, 177-198.

5. Feddersen, W. E., & Brown, W. L. Spatial learning from different positions in a room. *J. Genet. Psychol.*, 1956, **88**, 113-120

6. Gentry, G. V., Brown, W. L., & Kaplan, S. J. An experimental analysis of the spatial location hypothesis in learning. *J. Comp. & Physiol. Psychol.*, 1947, **40**, 309-322.

7. Gentry, G. V., Brown, W. L., & Lee, H. Spatial location in the learning of a multiple-*T* maze. *J. Comp. & Physiol. Psychol.*, 1948, **41**, 312-318.

8. Hebb, D. O. Organization of Behavior: A Neuropsychological Theory. New York: Wiley, 1949.

9. Honzik, C. H. The sensory basis of maze learning in rats. *Comp. Psychol. Monog.*, 1936, **13**, No. 4.

10. Hunter, W. S. The temporal maze and kinaesthetic sensory processes in the white rat. *Psychobiology*, 1920, **2**, 1-17.

11. Hunter, W. S. The sensory control of the maze habit in the white rat. *J. Genet. Psychol.*, 1929, **36**, 505-536.

12. Macfarlane, D. A. The rôle of kinesthesis in maze learning. *Univ. Calif. Pub. Psychol.*, 1930, **4**, 277-305.

13. Ritchie, B. F. Studies in spatial learning: VI. Place orientation and direction orientation. *J. Exper. Psychol.*, 1948, **38**, 659-669.

14. Small, W. S. An experimental study of the mental processes of the rat. *Amer. J. Psychol.*, 1900, **2**, 133-165.

15. Tolman, E. C., Ritchie, B. F., & Kalish, D. Studies in spatial learning: I. Orientation and the short-cut. *J. Exper. Psychol.*, 1946, **36**, 13-24.

16 ———. Studies in spatial learning: II. Place learning vs. response learning. *J. Exper. Psychol.*, 1946, **36**, 221-229.

17. Walton, A. Visual cues in maze running by the albino rat. *J. Genet. Psychol.*, 1930, **38**, 50-77.

18. Watson, J. B. Kinaesthetic and organic sensations: Their rôle in the reactions of the white rat. *Psychol. Rev. Monog.*, 1907, **8**, No. 2.

19. Weaver, H. E., & Stone, C. P. The relative ability of blind and normal rats in maze learning. *J. Genet. Psychol.*, 1928, **35**, 157-177.

*Department of Psychology*
*University of Texas*
*Austin 12, Texas*

*The Journal of Genetic Psychology*, 1957, **90**, 59-67.

# THE RÔLE OF MEDIATING VERBAL RESPONSES IN THE CONCEPTUAL SORTING BEHAVIOR OF NORMALS AND PARANOID SCHIZOPHRENICS*[1]

*Department of Psychology, University of Massachusetts*

James D. Fenn and Albert E. Goss

## A. Introduction

Gibson (4), Baum (1), and Osgood (8) have hypothesized that acquired equivalence of cues (internal generalization, common mediating responses) is a determinant of conceptual behavior in various concept formation situations. The complementary function of acquired distinctiveness of cues (2) in such situations, however, has not received equal attention. The theoretical analysis outlined here encompasses both mechanisms.

Hypothesized relationships among the block stimuli, mediating verbal processes, and sorting responses of a conceptual task similar to that described by Hanfmann and Kasinin (6) have been diagrammed in Figure 1. The 16 blocks are combinations of tall ($T$) and short ($Sh$), large ($L$) and small ($Sm$), top and bottom areas either circular ($C$) or square ($Sq$) in shape, and black ($B$) or white ($W$). It is assumed that instructions and/or prior learning result in associations between each of the four tall-large blocks and the responses, $R_{\text{tall-large}}$ or $R_{\text{lag}}$. Different common responses are evoked by the sets of four tall-small, short-large, and short-small blocks. The common mediating stimuli should enhance similarity of blocks within each category. Acquired distinctiveness based on different mediating stimuli for each category should decrease inter-category generalization based on color and shape. Therefore, these relationships of external and mediating stimuli should increase the probability of manipulative sorting of the blocks into height-size categories.

As a first step in testing this prediction, common familiar word or nonsense syllable responses to blocks within each category were learned. Then, a modification of the Hanfmann and Kasinin (6) procedure was used to ascertain whether there was an increase in the number of placements according to height-size.

---

*Received in the Editorial Office on September 2, 1955.

[1] A more complete account can be found in (3).

Various investigations (6) have reported differences in the conceptual sorting behavior of normals and schizophrenics. Therefore, the generality of the experimental hypothesis and possible interaction of experimental con-

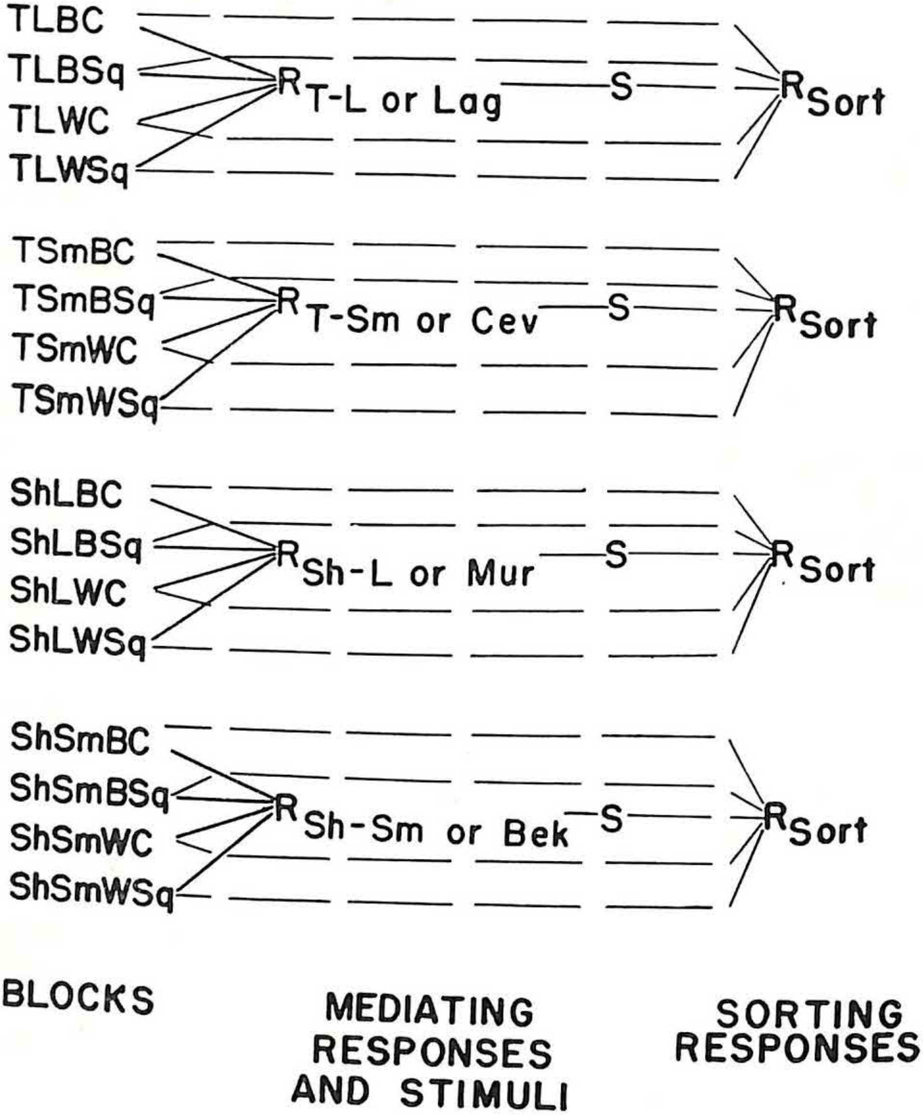

FIGURE 1
HYPOTHESIZED RELATIONSHIPS AMONG BLOCK STIMULI, MEDIATING RESPONSES AND STIMULI, AND SORTING RESPONSES OF A CONCEPTUAL SORTING TASK SIMILAR TO THAT DESCRIBED BY HANFMANN AND KASININ (6)

ditions with type of *S* were investigated by the use of both normals and paranoid schizophrenics.

## B. METHOD

### 1. *Subjects*

The *S*s were 40 normals (attendants, maintenance personnel, and nurses) and 40 non-deteriorated paranoid schizophrenics (33 from the Northampton State Hospital and 4 and 3 from the Worcester and Gardner State Hospitals, respectively). Both normals and paranoids were divided into four groups of five males and five females each. The resultant eight groups were equated for means of ages with a maximum range for any group of from 20 to 40 years.

*IQ* estimates for normals were obtained with the *CAS* short form of the Wechsler-Bellevue (Form I) which correlates +.93 with full scale values (9). Short form or full scale scores were used for the paranoids. While differences among means of *IQ*s of the four normal groups and within the four paranoid groups were not significant, the over-all mean of about 114 for the normals was significantly higher than that of 102 for the paranoids.

### 2. *Stimuli and Apparatus*

The 16 black or white, circular or square blocks were either 1 in. or ½ in. high and had top and bottom areas of 0.80 or 0.43 sq. in. Paired-associate procedures were used for acquisition of the familiar word or nonsense syllable responses.

The blocks were exposed one at a time by lifting the shutter of a 3 x 4 in. window in the front of a 9 in. high x 12½ in. wide x 6 in deep exposure-box. Raising the shutter activated a time-delay relay which turned on a light on the top of the box to end the 2-sec. anticipation interval.

There were eight different random orders of paired-associates, six of which were used also in presenting the blocks alone. Because of problems in holding the visual attention of paranoids, *E* told *S*s the correct label at the end of the anticipation period. On the average this required about 5 sec. Inter-block intervals averaged 5 sec. in length and the inter-trial intervals were about 10 sec.

A gray board 12 in. square and divided into four 6-in. squares by ½ in. wide black lines was added for the sorting task.

### 3. *Procedures for Block-Sorting Experiences*

Procedures for pre-sorting experiences and abbreviations for the different conditions and groups are presented in Table 1. Normals and paranoids learned the familiar word or nonsense syllable labels for blocks within each height-size category to a criterion of 14/16 correct anticipations.

TABLE 1
PRE-SORTING EXPERIENCES OF EACH GROUP OF 10 *S*s

| Pre-sorting experience or condition | Type of *S* | |
|---|---|---|
| | Normals (*NM*) | Paranoids (*P*) |
| Learn familiar words to 14/16 Criterion (*FW*)* | *FW-NM* | *FW-P* |
| Learn nonsense syllables to 14/16 criterion (*NS*) | *NS-NM* | *NS-P* |
| See, discriminate, name for six trials (*SDN*) | *SDN-Nm* | *SDN-P* |
| Control (*C*) | *C-Nm* | *C-P* |

*Abbreviations for conditions and type of *S* alone or in combination.

The see, discriminate, and name condition controlled for exposure to and naming experience with the blocks. Instructions for this condition were to pay attention to the blocks and to give them names which *S* considered appropriate in the 2-sec. interval before the light went on. The shutter remained up for about 5 sec. more to provide lengths of exposure to the blocks equal to those for the verbal learning groups. In order to avoid an *SDN* group for each verbal learning group, but at the same time provide an approximate match for amount of pre-sorting experience, means of total pre-sorting trials for the four verbal learning groups were averaged. The obtained value of 5.9 trials was rounded to the six trials administered to both *SDN-Nm* and *SDN-P*.

### 4. *Procedure for Block-Sorting Task*

Immediately after completion of pre-sorting experiences these groups were introduced to the block-sorting task. The controls undertook the task with no prior exposure to the blocks.

There were four trials. At the beginning of each trial one block from each of the four height-size categories was placed at the outside corner of one of the four squares of the sorting board. For the four trials, four different sets of four blocks, one for each category, were used. That is, each of the 16 blocks was on the board once. Placement of the four blocks of each set, one to each square, was random, subject to the restriction that a block for each category should appear once in each square.

Different random orders of the remaining 12 blocks of each trial were then presented in the window one at a time. As each block appeared, *S*s were instructed to place that block with the one corner block which they thought it matched. Thus, the blocks could presumably be sorted on the basis of height, size, shape, or color or any combination of these characteristics. There was no time limit on responses and no correction of placements was permitted. Inter-block and inter-trial intervals averaged 5 and 10 sec., respectively.

## C. RESULTS

### 1. *Pre-Sorting Experiences*

Because normals had higher *IQ*'s than paranoids, rho's were computed between *IQ*'s and both trials and errors to criterion for each verbal learning group. The obtained values of from —.44 to —.72 ($df = 8$) suggested the possibility of a significant relationship between intelligence and rate of learning. Accordingly, both analysis of variance and analysis of covariance were applied to the means of trials and errors to criterion (Table 2). Since both techniques led to the same conclusions, for the purpose of simplicity, only the results of the former will be presented.

TABLE 2

MEANS AND STANDARD DEVIATIONS OF TRIALS AND ERRORS TO CRITERION FOR THE VERBAL LEARNING GROUPS

| | | Type of *S* | | | |
|---|---|---|---|---|---|
| | | *Nm* | | *P* | |
| Measure | Conditions | Mean | *SD* | Mean | *SD* |
| Trials to criterion | *FW* | 1.5 | 0.7 | 3.7 | 2.4 |
| | *NS* | 4.2 | 2.2 | 10.3 | 2.9 |
| Errors to criterion | *FW* | 13.2 | 4.4 | 34.7 | 20.5 |
| | *NS* | 38.6 | 18.7 | 96.7 | 27.9 |

The significant *F*'s for main effects and interaction for both response measures (Table 3) when coupled with *t*'s for pairs of means indicated that: (*a*) normals learned familiar words and nonsense syllables more rapidly than comparable paranoid groups, (*b*) *NS-Nm* mastered nonsense syllables in approximately the same number of trials required for *FW-P* to reach criterion, and (*c*) the advantage of familiar words over nonsense syllables was relatively less for normals than for paranoids.

### 2. *Block-Sorting*

Because there were no marked changes in number of height-size placements over the four block-sorting trials only means of such placements for all four trials have been summarized (Table 4). Correlation of *IQ*s and height-size placements of each group yielded non-significant rho's of from —.21 to +.52 (*df* = 8). Since sex had no differential effect, differences among means were evaluated by a double-classification analysis of variance (Table 5). The significant *F* of 32.49 for conditions indicated that familiar word and nonsense syllable experiences led to marked increases in the number of height-size placements. However, the main effect of type of *S* was not sig-

TABLE 3
ANALYSIS OF VARIANCE OF TRIALS AND ERRORS TO CRITERION

| Measure | Source | *df* | Mean square | *F* |
|---|---|---|---|---|
| Trials | Conditions (*Cn*) | 1 | 170.23 | 32.18* |
| | Subjects (*S*s) | 1 | 216.23 | 40.88* |
| | *Cn* x *S*s | 1 | 40.02 | 7.56* |
| | Within | 36 | 5.29 | |
| Errors | *Cn* | 1 | 19,096.90 | 44.09* |
| | *S*s | 1 | 15,840.40 | 36.57* |
| | *Cn* x *S*s | 1 | 3,348.90 | 7.73* |
| | Within | 36 | 433.17 | |

*Significant at or beyond the 1 per cent level.

TABLE 4
MEANS AND STANDARD DEVIATIONS OF NUMBER OF CORRECT HEIGHT-SIZE PLACEMENTS OVER FOUR BLOCK-SORTING TRIALS

| | Type of *S* | | | |
|---|---|---|---|---|
| | *Nm* | | *P* | |
| Conditions | Mean | *SD* | Mean | *SD* |
| *FW* | 37.8 | 11.3 | 32.6 | 13.7 |
| *NS* | 38.0 | 10.3 | 25.9 | 20.4 |
| *SDN* | 5.3 | 7.2 | 6.4 | 6.6 |
| *C* | 6.5 | 5.7 | 6.9 | 6.8 |

TABLE 5
ANALYSIS OF VARIANCE OF HEIGHT-SIZE PLACEMENTS

| Source | *df* | Mean square | *F* |
|---|---|---|---|
| *Cn* | 3 | 5,006.22 | 32.49* |
| *S*s | 1 | 312.05 | 2.02 |
| *Cn* x *S*s | 3 | 189.35 | 1.22 |
| Within | 72 | 154.07 | |

*Significant at beyond the 1 per cent level.

nificant. With the exception of the *t*'s for differences between the two normal groups and *NS-P*, none of the differences between pairs of means of verbal learning groups and between pairs of means of the four remaining groups was significant at the 5 per cent level.

Analysis of block-placements of the *SDN* and *C* groups indicated that, in all four groups, blocks were matched according to shape, color, height, and size in that order. With the exception of a single inversion in one group, the order of decreasing preferences for second-order categories was shape-color, shape-height, shape-size, color-height, color-size, and height-size.

## D. DISCUSSION

### 1. *Pre-Sorting Experiences and Retention of Verbal Responses*

The fact that familiar (and more meaningful) words were learned more rapidly than nonsense syllables by both normals and paranoids, was consistent with other data (8, 10). Also, the superiority of normals to paranoids in learning the same verbal responses confirms and extends other observations of the relatively slower learning of schizophrenics (7).[2]

### 2. *Block-Sorting*

As predicted on the basis of the relationships among external and mediating stimuli in Figure 1, acquisition of common familiar word or nonsense syllable responses to all blocks within each of four height-size categories occasioned a significant increase in manipulative sorting according to height-size. Furthermore, analysis of the placements of the *SDN* and *C* groups suggested that this increase was against initial preferences for sorting in terms of shape and color both separately and in combination.

When learned to criterion, nonsense syllables were as effective in increasing height-size placements as familiar words. It is possible, however, that acquisition of nonsense syllable responses is a more complex process which involves the more or less simultaneous strengthening of associations between blocks and familiar words, between familiar words and nonsense syllables, and between blocks and nonsense syllables. The see, discriminate, and name experience did not affect height-size placements for either normals or paranoids.

Although there was a tendency for verbal learning normals to make more height-size placements than the paranoids, only the differences involving

---

[2] While paranoids learned better than other schizophrenics only the most coöperative of this group approached the performance level of the normals (7).

*NS-P* were statistically significant. It is possible that somewhat larger differences might have been obtained with older and/or more deteriorated paranoids or with *S*s in other clinical groups.

The relative equality of the sorting behavior of the verbal learning normals and paranoids could be interpreted as due to the strengthening of the *abstract set* (5) of the paranoids. Such a translation of *abstract set* into acquired equivalence and acquired distinctiveness has the advantages of incorporating this concept within a more precise theoretical context and of introducing the possibility of its manipulation through variations in the strength of mediating responses.

## E. Summary

The mechanisms of acquired equivalence and acquired distinctiveness of cues were applied to a conceptual block-sorting situation to generate a prediction that common familiar word or nonsense syllable responses to blocks within tall-large, tall-small, short-large, and short-small categories would increase manipulative placements according to height-size. The stimuli were 16 blocks representing the combinations of tall or short, large or small and circular or square top and bottom areas, and black or white. Groups of normals and paranoid schizophrenics learned familiar word or nonsense syllable responses to blocks within the four height-size categories by the paired-associate method to 14/16 correct anticipations. An additional group of normals and one of paranoids saw the blocks under instructions to see, discriminate among, and name them. There were 10 *S*s in each of these six groups and in normal and paranoid control groups.

In general, learning of verbal responses was more rapid for familiar words and by normals. However, differences between normals with nonsense syllables and paranoids with familiar words were not significant. In accordance with the hypothesis, marked increases in height-size placements were observed for both normals and schizophrenics who had learned the common familiar word or nonsense syllable responses to blocks within the height-size categories. Differences among the verbal learning groups, however, were not statistically significant.

## References

1. Baum, M. H. A study in concept attainment and verbal learning. Unpublished Ph.D. dissertation, Yale Univ., 1951.
2. Dollard, J., & Miller, N. E. Personality and Psychotherapy. New York: McGraw-Hill, 1950.

3. FENN, J. D. The rôle of mediating verbal responses in the conceptual sorting behavior of normals and paranoid schizophrenics. Unpublished M.S. thesis, Univ. Massachusetts, 1954.
4. GIBSON, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196-229.
5. GOLDSTEIN, K., & SHEERER, M. Abstract and concrete behavior; an experimental study with special tests. *Psychol. Monog.*, 1941, No. 239.
6. HANFMANN, E., & KASININ, J. Conceptual Thinking in Schizophrenia. New York: Nervous and Mental Diseases Monog., 1942.
7. HUSTON, P. E., & SHAKOW, D. Learning capacity in schizophrenia. *Am. J. Psychiat.*, 1949, **105**, 881-888.
8. OSGOOD, C. E. Method and Theory in Experimental Psychology. New York: Oxford Univ. Press, 1953.
9. RABIN, A. I. A short form of the Wechsler-Bellevue Test. *J. Appl. Psychol.*, 1943, **27**, 320-324.
10. UNDERWOOD, B. J. Experimental Psychology. New York: Appleton-Century-Crofts, 1949.

*Department of Psychology*
*University of Massachusetts*
*Amherst, Massachusetts*

*The Journal of Genetic Psychology*, 1957, **90**, 69-74.

# THE RÔLE OF MEDIATING VERBAL RESPONSES IN THE CONCEPTUAL SORTING BEHAVIOR OF CHILDREN*

*Department of Psychology, University of Massachusetts*

---

JANICE E. CAREY AND ALBERT E. GOSS

---

## A. INTRODUCTION

Fenn and Goss (1) observed that acquisition of common verbal responses to blocks within height-size categories increased subsequent height-size placements by normal and paranoid schizophrenic adults. In order to obtain some picture of the interaction of Fenn and Goss' conditions and amount of general experience in learning and using labels, a replication of their conditions with relatively less sophisticated preschool age children seemed desirable.

Previous experience with verbal learning of young children (5) and their performance in preliminary experimentation, combined with limitations imposed by school routine soon indicated, however, that exact replication of stimuli, procedures, and hypothesis would not be feasible. Therefore, the number of dimensions was reduced from 4 to 3 and, concomitantly, the number of block stimuli from 16 to 8. Instead of instructions which merely called for matching each presented block with one of four already on the sorting board, the children were reinforced for each correct height-size placement. This shift from *matching* to *learning* entailed a change in the experimental hypothesis to the prediction that acquisition of common labels for blocks within each height-size category would occasion more rapid learning to match blocks according to height-size.

## B. METHOD

### 1 *Subjects*

Forty 4- to 5-year-old children from nearby nursery schools and a kindergarten were divided into four groups of five boys and five girls each.[1] While not completely successful, an attempt was made to assign boys or girls

---

*Received in the Editorial Office on September 2, 1955.

[1] The authors are indebted to Mrs. W. H. Thies (University of Massachusetts Nursery School), Mrs. L. L. Derby (Derby Day School), Mrs. L. R. Grose (Grose Nursery School), and Misses Evelyn Beyer and Helen Rees (Smith Nursery and Day Schools) for their assistance in providing *S*s and experimental facilities.

from the same school and of approximately the same age in months to all groups. Mean ages of the groups were from 4-6 to 4-7. In general, the parental background was college-educated, and business or professional (often college teaching).

### 2. *Stimulus Materials*

The eight white of Fenn and Goss' (1) 16 blocks were eliminated to leave four circular and four square black blocks which were tall-large, tall-small, short-large, or short-small. *Lag*, *cev*, *mur*, and *bik* were assigned to these four height-size categories, respectively. Preliminary experimentation, however, indicated that high-big, high-little, low-big, and low-little would be more suitable than the familiar words used previously. Eight different random orders of eight block-label pairs or of blocks alone were prepared.

The exposure box and sorting-board were those used by Fenn and Goss (1). For the paired-associate learning the blocks were presented one at a time with a lengthened anticipation interval of 5 sec. after which *E* gave and *S* repeated the correct label for the block. This required about 5 more sec.

### 3. *Pre-Sorting Experiences*

The three pre-sorting experiences were learning familiar word (*FW*) or nonsense syllable (*NS*) labels for the blocks and a see, discriminate and name condition. Paired-associate learning of the familiar word or nonsense syllable responses to both blocks of each height-size category continued to criteria of ⅞ correct anticipations or 12 trials. The latter criterion was introduced because preliminary experiences had indicated that continuation beyond this point would be handicapped by *S*s' boredom and interference with school activities. Also, to activate and hold the children's attention, they were told that they would be given balloons after they were able to name some of the blocks. All *S*s were given three balloons; for most, one was given after Trials 4, 8, and 12. To allow for shifts in *S*s' attention inter-block intervals were not controlled. They averaged about 5 sec. The inter-trial intervals were about 10 sec.

Exposure to and naming experiences with the blocks was controlled by the see, discriminate, and name (*SDN*) condition in which *S*s were instructed to look at each block and to guess its name. Each block alone was exposed for about 10 sec. with inter-block and inter-trial intervals which averaged 5 and 10 sec., respectively. Every *S* of *SDN* had 12 trials.

### 4. *Block-Sorting*

After a 2-min. break to minimize boredom the pre-scoring experience groups were introduced to the sorting task. The controls had no prior experience with the blocks.

The task was to learn to sort the blocks into four height-size categories. At the beginning of each trial *E* placed one of the two blocks of each category at the outer corner of the four squares of the sorting-board. Different combinations of one block for each category were used for each of the 10 trials with random assignment to the corners. The remaining four blocks of each trial were presented randomly one at a time and *S*s were instructed to put each in the corner containing the block with which *S* thought the presented block belonged. Correct placements were reinforced by *E* saying "right." Nothing was said after incorrect placements. There were about 30 sec. between successive trials.

## C. RESULTS

### 1. *Verbal Learning and Retention*

Only two *S*s in each verbal learning condition reached the ⅞ criterion. Although group *FW* was consistently superior to *NS*, *t*'s for differences between means of correct responses for all trials and for the last trial of the ⅞ or 12 trial criteria were not significant (Table 1).

TABLE 1
MEANS AND STANDARD DEVIATIONS OF CORRECT VERBAL RESPONSES FOR ALL LEARNING TRIALS AND FOR THE LAST LEARNING TRIAL

| Group | All trials Mean | All trials *SD* | Last trial Mean | Last trial *SD* |
|---|---|---|---|---|
| *FW* | 31.4 | 9.2 | 5.0 | 1.9 |
| *NS* | 23.7 | 12.3 | 3.5 | 2.1 |
| *t** | 1.51 | | 1.58 | |

*A *t* of 2.10 is necessary for significance at the 5 per cent level with 18 *df*.

### 2. *Block-Sorting*

The mean of 22.6 correct placements for all 10 block-sorting trials for Group *FW* was higher than the means for the other conditions (Table 2). The *F* of 4.59 for 3 and 36 *df* indicated that the superiority of this group could not be readily attributed to chance. This conclusion is supported by the plot of correct responses as a function of two-trial blocks (Figure 1).

While *S*s who had learned familiar words progressed from an initial chance level to means of 6.2 and 6.1 correct responses for Trials 7-8 and 9-10 respectively, the other groups continued to respond with chance frequencies.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF CORRECT PLACEMENTS FOR ALL 10 TRIALS

| Group | Mean | *SD* |
|---|---|---|
| *FW* | 22.6 | 6.1 |
| *NS* | 11.4 | 5.3 |
| *SDN* | 13.9 | 9.0 |
| *C* | 13.9 | 6.5 |

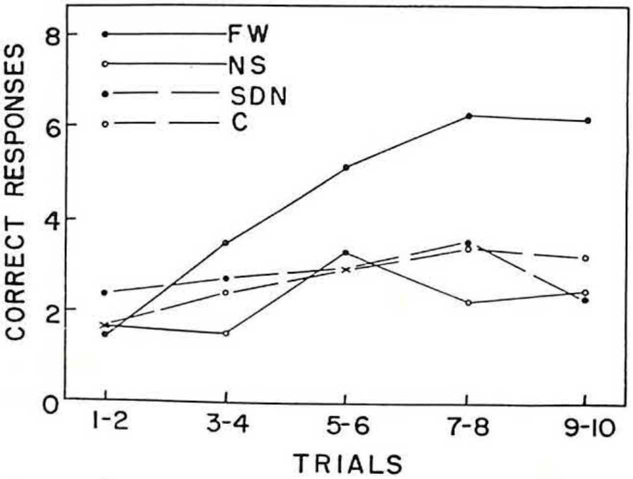

FIGURE 1
MEANS OF CORRECT RESPONSES FOR SUCCESSIVE TWO-TRIAL BLOCKS

## D. DISCUSSION

### 1. *Verbal Learning*

The contrast between the rates of verbal learning of these children and Fenn and Goss (1) adults is striking. Adults reached a 14/16 criterion in 1.5 and 3.7 trials with familiar words and 4.2 and 10.3 trials with nonsense syllables. With half as many stimuli and a lengthened anticipation interval, 16 of 20 children failed to reach the 7/8 criterion in 12 trials. That this relatively slow learning is not unique to block stimuli is indicated by a similar rate for children's acquisition of discriminative labels for four white squares of different areas (5).

The failure to obtain significantly more rapid learning for Group *FW* than for nonsense syllable *S*s may have been due to curtailment in the extent of differences between means of trials and correct responses to a 7/8 criterion imposed by the 12-trial maximum. Also, since *S*s in the two groups were only making 40 and 60 per cent correct responses on the last trial the point of greatest difference between means of correct responses on a single trial might not have been reached.

### 2. *Block-Sorting*

Acquisition of familiar words or nonsense syllables increased the height-size placements of adults without reinforcement of correct responses (1). The first height-size placements of the children who had learned verbal responses were at chance levels. With continued reinforcement Group *FW* but not Group *NS* exhibited the predicted more rapid acquisition of height-size placements.

While learning nonsense syllables was not statistically inferior to acquisition of familiar words this difference may have been a practical one if some minimum level of mastery of verbal responses is necessary for measurable positive transfer to subsequent discriminative tasks (2). Also, familiar words may have had the advantage of greater acquired distinctiveness based on larger numbers of associated responses (2, 3, 4).

Much greater mastery of verbal responses than that attained by Group *FW* is probably a minimum condition to obtain the increased height-size placements without reinforcement. Furthermore, both as stimuli and responses, familiar words and nonsense syllables were probably more meaningful and better learned for adults and thus contributed to their superior verbal learning and conceptual sorting behavior (3, 4).

It should be noted that, although there was random variation of blocks and their board positions, the introduction of reinforcement could be considered as transforming conceptual sorting into a complex discrimination learning task. The possibility of drawing such a distinction in clear-cut fashion, however, seems remote.

## E. SUMMARY

Preschool age children were divided into four groups of 10 *S*s each equated for sex, age, and school. Eight blocks representing combinations of tall or short, large or small, and circular or square were presented one at a time in the window of box-like exposure device.

One group learned common familiar word and another common nonsense syllable labels for the two blocks within each height-size category to 7/8 correct anticipations or for 12 trials. A third group saw and named the blocks for 12 trials. These groups plus controls then had 10 trials to learn to sort the blocks according to height-size.

The experimental hypothesis of positive transfer from verbal learning to learning to sort was confirmed by the familiar word but not by the nonsense syllable group. Reason for this outcome and the bases of the children's marked inferiority to adults in both verbal learning and conceptual sorting were considered.

## References

1. FENN, J. D., & GOSS, A. E. The rôle of mediating verbal responses in the conceptual sorting behavior of normals and paranoid schizophrenics. *J. Genet. Psychol.*, 1957, **90**, —.
2. GOSS, A. E. A stimulus-response analysis of the interaction of cue-producing and instrumental responses. *Psychol. Rev.*, 1955, **62**, 20-31.
3. NOBLE, C. E. The meaning-familiarity relationship. *Psychol. Rev.*, 1952, **59**, 421-430.
4. SHEFFIELD, F. D. The rôle of meaningfulness of stimulus and response in verbal learning. Unpublished Ph.D. dissertation, Yale Univ., 1946.
5. SMITH, S. L., & GOSS, A. E. The rôle of the acquired distinctiveness of cues in the acquisition of a motor skill in children. *J. Genet. Psychol.*, 1955, **87**, 11-24.

*Department of Psychology*
*University of Massachusetts*
*Amherst, Massachusetts*

*The Journal of Genetic Psychology*, 1957, **90**, 75-82.

# AN EVALUATION OF SHOCK'S "GROWTH CURVES"*

*Haverford College and University of British Columbia*

ROBERT J. WISNER AND DOUGLAS T. KENNY

## A. INTRODUCTION

In the current psychological literature, the formulation of rational equations for growth has been a neglected research area. One of the most recent attempts to develop a mathematical formulation for describing and predicting growth is that provided by N. W. Shock in the *Handbook of Experimental Psychology* (1). Since Shock's paper appears as a chapter in such a prominent book and is one of the few published general discussions of growth, his work merits serious consideration. While his paper opens the horizons to a potentially fruitful field of investigation and contains some interesting new derivations of important, well-known growth models, the present authors wish to take issue with certain fundamental errors of conception and omission found in his quantitative treatment of growth. Hence the purpose of this paper is to: (*a*) challenge and to show the limitations of several of the statements and mathematical derivations contained in the work; (*b*) make appropriate corrections to and expansions of the mathematical formulations; (*c*) provide suitable interpretations for the parameters (where possible); and (*d*) note certain typographical errors. More explicitly, the procedure will be to present and then to evaluate critically Shock's basic growth equations and four cases of his general theory.

It is our hope that these criticisms will clarify Shock's article, and that the corrections and formula generalizations will prove useful to those interested in growth experimentation.

## B. SHOCK'S FUNDAMENTAL EQUATION OF GROWTH

A brief review of Shock's formal approach to growth curves will prove useful in order to clarify certain basic distinctions we will make. Using his notation with slight modifications, let

$N(t)$ = the total number of living cells at time $t$,

$\Phi(N,t)$ = the rate of change of $N(t)$ due to intrinsic cellular factors, and

$\Theta(N,t)$ = the rate of change of $N(t)$ due to the interaction between cells.

---

*Received in the Editorial Office on November 1, 1955.

Assuming that the rate of change in N(t) is the sum of the rates of change $\Phi$ and $\Theta$, Shock begins with

$$\frac{dN(t)}{dt} = \Phi(N,t) + \Theta(N,t) \tag{1}$$

Concentrating on the first summand of (1), he assumes that it is plausible to suppose that $\Phi(N,t)$ is proportional to N(t). Allowing his "constant of proportionality" to be a function of N(t) and *t*, he arrives at

$$\Phi(N,t) = A(N,t) \cdot N(t) \tag{2}$$

He then analyzes $\Theta(N,t)$ and observes that it must depend upon N(t) but that this dependence need not be linear. The formulation is

$$\Theta(N,t) = B(N,t) \cdot [N(t)]^k \tag{3}$$

and he implies that, in general, the exponent *k* need not be constant.

Substituting (2) and (3) in (1), Shock arrives at what he calls "the fundamental equation of growth":

$$\frac{dN(t)}{dt} = A(N,t) \cdot N(t) + B(N,t) \cdot [N(t)]^k \tag{4}$$

Shock contends that ". . . all the important rational growth equations . . . can be derived by making logical assumptions with regard to the functions *A* and *B*."

## C. CRITICAL ANALYSIS OF THE EQUATION OF GROWTH

It is not proposed here to take issue with the assumptions Shock makes for one may, of course, make whatever assumptions one chooses, provided that the assumptions are consistent and that the conclusions are logically drawn. But there are several difficulties in what Shock claims follows from these assumptions. First, if $\Phi(N,t)$ is proportional to N(t) as is postulated, one would expect to find equation (2) reading $\Phi(N,t) = A \cdot N(t)$, where *A* is a constant—this being what is usually meant by "proportional." As it stands, equation (2) says nothing, for the right-hand side of that equation is a perfectly general function of *N* and *t*. Next, the analysis of $\Theta(N,t)$ does not follow from any stated or previously implied assumption. That the dependence of $\Theta(N,t)$ upon N(t) need not be linear is not the basis for equation (3) any more than it would be the basis for writing $\Theta(N,t) = \sin N$. Moreover, the same criticism of equation (3) is made as above for equation (2), i.e., the right-hand side of equation (3) is a perfectly general function of *N* and *t*. As written, then, Shock has failed to analyze or to specify either summand of equation (1); hence, equation (4) is not more specific than is equation (1). Thus, to call equation (4) the "fundamental equation of growth" would be misleading unless it were really true that all the important rational growth equations were derivable from it. That this is not the case, however, can be realized by recalling the important *random net model* for growth which pervades the literature of the Rashevsky group. It is easily seen that the *random net model* is not derivable from equation (4) by any choice of A(N,t) and B(N,t)—the reason being that "time" in the random net is not a continuous variable, and so derivatives with respect to it are impossible.

Finally, in the four special cases of the "general theory" which Shock presents, he does not follow his suggested procedure of making logical assumptions about A(N,t) and B(N,t). In each case, he refers directly to equation (1) and analyzes $\Phi(N,t)$ and $\Theta(N,t)$, a procedure which the present authors heartily endorse.

## D. Application of the Growth Equation

Following his development of the fundamental equation of growth, Shock examines some special cases of the general theory and discusses four growth curves. Each of the four models will be presented here essentially as they appear in Shock's article with criticisms, expansions and corrections following.

### 1. *Case I*

In analyzing $\Phi$ it is assumed that the cells have equal and constant division times, thereby yielding

$$\Phi = AN \tag{5}$$

"where *A* is a true constant (with transient or lag effects)" (1, p. 332).

In analyzing $\Theta$ it is assumed that each cell produces a substance harmful to every cell (including itself), and that the "proportionality constant is *B*." He deduces that

$$\Theta = -BN^2 \tag{6}$$

since "the total number of effects of *N* cells upon themselves is BN . . . (and) . . . the number of effects of cells upon their fellows is BN(N — 1)" making a total of BN + BN(N —1) = BN². Hence

$$\frac{dN}{dt} = AN - BN^2 \tag{7}$$

Equation (7) is purported to integrate directly into

$$N(t) = \frac{A/B}{1 + (A/B - 1)e^{-At}} \tag{8}$$

In examining this case, one notices that the value of $\Theta$ is not just the result of interaction between cells but includes the effect of each cell upon itself, and of course this should be included in the derivation of $\Phi$. It is easy to see that the correct values should be

$$\Phi = AN - BN \tag{9}$$

$$\Theta = - BN(N - 1) \tag{10}$$

This latter analysis gives the same result for dN/dt.

Shock's statement that *A* is a "true constant (with transient or lag effects)" is obviously a typographical error. Reference to the Shock and Morales article (2) indicates that *A* is meant to be a true constant *without* transient or lag effects.

A rather serious error of omission is committed by Shock when *A* and *B* are not interpreted in terms of growth, for otherwise they remain parameters to be determined by some curve fitting process. We present here what we believe to be the correct interpretation: *A* equals the number of cells into which each cell divides per unit of time, and *B* is the probability that in a unit of time the substance produced by each cell will harm some cell. "Harm" here is obviously taken to mean *destroy* or *kill*. Another interpretation of "harm" might be that the cell is rendered unreproductive. Such an interpretation with its corresponding model, which is easily derivable, might prove useful in some experiments.

Finally, equation (7) "integrates directly" into (8) only if one assumes that N(O) = 1 or an equivalent statement. The proper solution of (7) which allows the experiment to begin with any number N(O) of initial cells is

$$N(t) = \frac{A/B}{1 + (A/BN(O) - 1)e^{-At}} \tag{11}$$

## 2. *Case II*

The next model makes the same assumptions that are made in Case I, except that "the proportionality constant *B* changes in time." *B* is expressed as a polynomial in *t* to obtain

$$\Theta = -N^2 \sum_{s=o}^{s=n} K_s t^s \tag{12}$$

Substituting (12) and (5) in (1) and integrating, Shock obtains

$$N(t) = \frac{1}{Ke^{-At} + e^{-At}\int B(t)e^{At}dt} \qquad (13)$$

where "*K* is a constant of integration."

There seems to be confusion about the concept of a constant when dealing with *B*. Constants, of course, do not "change in time." However, even assuming that *B* is a function of time, there is no reason for expressing *B* as a polynomial in *t*, since clearly not all functions are polynomials. Even if a basis could be shown for writing *B* as a polynomial, then some indications of the degree of the polynomial and growth interpretations of its parameters should have been expressed. Moreover, the present authors would not consider the fact that every continuous function can be approximated as closely as you please on a finite interval by a polynomial as a sufficient "basis." The reason is that the constants are not then interpretable as growth parameters. Additionally, the reason that (13) does not utilize the expression of *B* as a polynomial is not clear. More specifically, equation (13) should read

$$N(t) = \frac{1}{Ke^{-At} + \sum_{s=0}^{s=n} (-1)^s A^{-s-1} B^{(s)}(t)} \qquad (14)$$

where

$$B^{(s)}(t) = \frac{d^s B(t)}{dt^s}$$

and

$$K = \frac{1}{N(O)} + \sum_{s=0}^{s=n} (-1)^{s+1} A^{-s-1} K_s$$

### 3. *Case III*

In this case the stability of the population is assumed to be affected only by retarding interaction effects which are proportional to the population. Hence, for positive constants *A* and *B*, $\Phi = A$ and $\Theta = -BN$ yielding $dN/dt = A - BN$, whence

$$N(t) = \frac{A}{B}(1 - e^{-Bt}) \qquad (15)$$

It should be noted that (15) is correct only if one assumes that $N(O) = O$. Again, we give the formula correct for any number of initial cells:

$$N(t) = \frac{A}{B} + [N(O) - \frac{A}{B}]e^{-Bt} \qquad (16)$$

### 4. *Case IV*

In this model, it is assumed that there is always an amount of growth stimulant proportional to the number of cells. His analysis, by the assumption of Fechner's law, is

$$\Phi = (-C \log A)\ N \qquad (17)$$

and

$$\Theta = CN \log N \qquad (18)$$

Substituting (17) and (18) in (1), he obtains

$$\frac{dN}{dt} = (\log N - \log A)CN \qquad (19)$$

and hence

$$N(t) = Ae^{Be^{Ct}} \qquad (20)$$

where $B$ is a constant of integration.

The main difficulty here is that equation (20) is not plausible, for let us calculate $N(O)$:

$$N(O) = Ae^{B}$$

Hence

$$e^{B} = N(O)/A$$

and

$$N(t) = A[N(O)/A]^{e^{Ct}} \qquad (21)$$

Thus, if $N(O)/A > 1$, the number of cells grows with time, and if $N(O)/A < 1$, the population diminishes and eventually dies out, and finally, if $N(O)/A = 1$, the number of cells remains constant. Since $A$ is essentially a measure of the amount of growth stimulant produced by each cell, one would expect that an increased $A$ would increase growth, and conversely. It is unreasonable that $N(O)$ should have to be large relative to $A$ for growth.

Of course, the difficulty is easy to analyze. If we assume, as Shock does, that the amount of stimulating substance present at any time is $AN$, and in addition, if we assume Fechner's law, then the effect on each cell is $C \log AN$, where $C$ is constant. Hence, the effect on $N$ cells is $CN \log AN$, so

$$\frac{dN}{dt} = CN \log AN = CN(\log A + \log N) \qquad (22)$$

The solution of (22), with the usual boundary conditions, is

$$N(t) = \frac{1}{A}\,[AN(O)]^{e^{Ct}} \qquad (23)$$

It is clear that the error in Shock's analysis is with the use of the minus sign in (17). There is also an error parallel to one made in Case I in that $\Theta$ is not due completely to interactive effects. The corrected model has the following properties: If $AN(O) > 1$, the body grows; if $AN(O) < 1$, the body eventually dies; and if $AN(O) = 1$, then $N(t) = 1/A$ for all $t$, producing a stable population of cells. Now if the amount of stimulating substance is increased, growth increases—as one would expect.

It might be mentioned in passing that minor additional errors, possibly typographical, we noticed. His equation (20) should read

$$\frac{dM}{dt} = \frac{M}{N}\,\frac{dN}{dt} + \frac{M}{p}\,\frac{dp}{dt} + \frac{M}{V}\,\frac{dV}{dt}$$

And his equations (21) through (24), besides containing the errors noted previously, are correct only to within an additive constant.

## E. General Remarks

Shock claims that he has rendered the well-known Gompertz curve, equation (20), rational, remarking that it has "up to now been regarded as purely empirical." It is very interesting that the Gompertz curve is derivable by use of Fechner's law, but, as is well known, Fechner's law is empirical; and though the Gompertz curve is derivable "rationally" from Fechner's law, its basis remains empirical. Thus, the present authors at least will continue to consider the Gompertz curve empirical.

Shock then explains that "Since there is a many-to-one correspondence between an integral form and a differential equation, the interpretation of the constants . . . is not unique" (1, p. 334). (Parenthetically, it should be added that the constants referred to in this quote are evidently constants of integration, since it is through these that such a many-to-one correspondence exists.) We deem it important here to explain what are called "boundary value" conditions in differential equations. For each constant of integration present in the integral form, one needs to know but one set of values for all variables present; and hence, for the models discussed in this paper, one needs to know only the number of cells at any particular time. Under most experimental conditions, one would know (or at least have a good estimate of) the number of cells he begins with, and so we have assumed that N(O) is known and have evaluated the constants of integration in terms of this. There is no ambiguity about the constants of integration for any experimenter using these models if he knows any such fact about the experiment.

Finally, one would have hoped to find in Shock's article some application of the models suggested by him. Unfortunately, in his last section on the types of growth curves found in actual experimental investigation, he makes no attempt to apply these models to the experimental growth curves.

## REFERENCES

1. SHOCK, N. W. Growth curves. In (S. S. Stevens, *Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951. (Pp. 330-346.)
2. SHOCK, N. W., & MORALES, M. F. A fundamental form for the differential equation of colonial and organism growth. *Bull. Math. Biophy.*, 1942, **4**, 63-71.

*Department of Mathematics*
*Haverford College*
*Haverford, Pennsylvania*

*Department of Psychology*
*University of British Columbia*
*Vancouver, B. C., Canada*

*The Journal of Genetic Psychology*, 1957, **90**, 83-88.

# ADAPTIVE RESPONSES AND THE CONTIGUITY PRINCIPLE*[1]

*Department of Psychology, Cornell University*

GEORGE W. BOGUSLAVSKY

## A. PROBLEM

It is generally recognized that unconditioned and conditioned responses are not exact duplicates of each other, the latter having been described as preparatory or adaptive in nature (2, p. 39). This qualitative discrepancy has led some psychologists to reject association by contiguity as a valid principle of learning (3, p. 20). One of the suggested alternatives is to regard *CR* as an operant form of behavior, in which case no attempt is made to specify the stimuli (5). Another alternative is to view the *CR* as a consequence of some central process, which raises the problem of defining the relation between the central process and the reaction (6).

The rejection of contiguity is understandable, since behavior which ultimately emerges cannot be traced directly to the effect of the unconditioned stimulus. It is also inescapable, if the molar approach to the study of behavior restricts observation to the initial stimuli and final outcomes. In contrast to such molar treatment, observations described in this paper are not restricted to the initial and terminal variables, but include progressive changes in behavior which culminate in the adaptive response. The aim of this study was to examine, in a manner consistent with *S-R* theory, the possibility of identifying the factors which lead to the *CR* in its final form.

## B. METHOD

The observations were made at the Cornell Behavior Farm on nine goats. The animals, placed in a light restraining harness, were presented with a *CS* (dimming of illumination or a buzzer) for 2 to 10 seconds, followed by a 6 to 10 v. electric shock to the right foreleg, of ½ sec. duration. Six of the animals were shocked only if the leg was resting on the platform at the end of the signal; the remaining three were shocked on every trial. The foreleg movements of seven animals were recorded in detailed written protocols, and of the remaining two on polygraph tape.

---

*Received in the Editorial Office on November 2, 1955.

[1] Read at the 1955 meetings of the Eastern Psychological Association.

## C. Results

The results indicate that progress of the formation of an adaptive avoidance reaction may be divided into three stages: (*a*) initial stage, during which *CR* consists of a single flexion of the limb, occurring shortly after the onset of *CS* and practically indistinguishable in form from *UR*; this observation confirms an earlier finding by others (1); (*b*) intermediate stage, during which *CR* consists of a series of rapid flexions, with the limb contacting the platform after each flexion; and (*c*) final stage, during which *CR* consists of a series of alternate flexions and extensions, without the limb touching the platform. In an avoidance conditioning experiment this constitutes an adaptive reaction. The results are shown graphically in Figures 1 through 4.

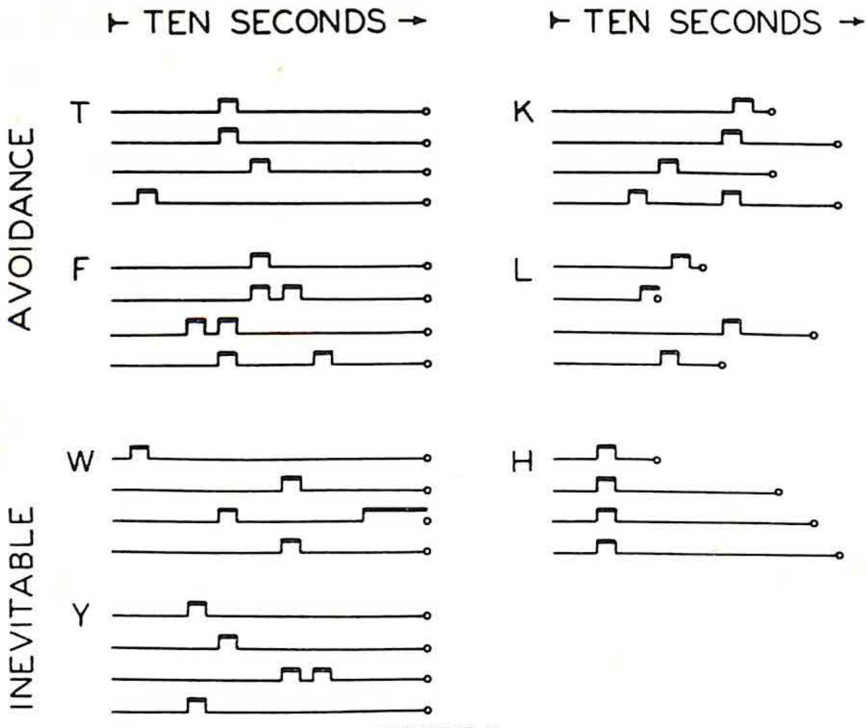

FIGURE 1
Incidence and Duration of Flexions During the First 4 *CR's*
(Letters at the left identify individual animals. Small circles indicate termination of *CS*.)

Figure 1 shows the incidence and duration of leg retractions during the first four *CR*'s of seven animals. The values are approximate, since the information was originally recorded in written protocols with the aid of a metronome. The initial stage, indicated by a single elevated line, may last from one to several *CR*'s; one phenomenon, however, to which there has been no exception, is that the first *CR* always consists of a single flexion.

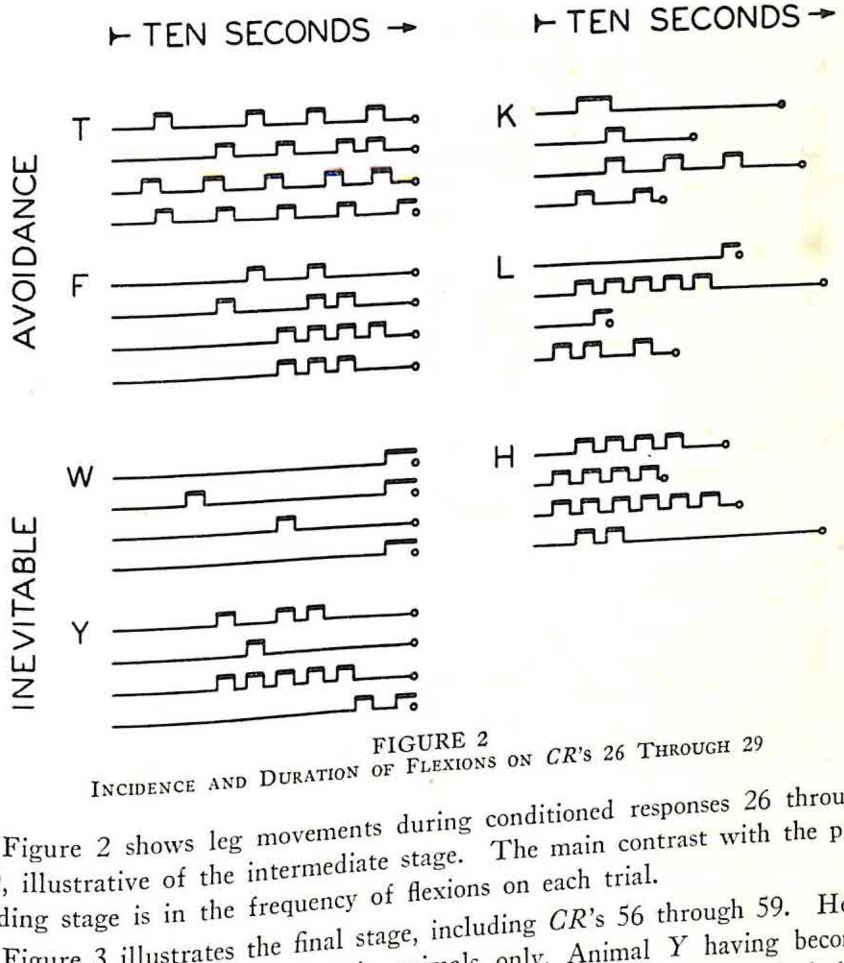

FIGURE 2

Incidence and Duration of Flexions on *CR*'s 26 Through 29

Figure 2 shows leg movements during conditioned responses 26 through 29, illustrative of the intermediate stage. The main contrast with the preceding stage is in the frequency of flexions on each trial.

Figure 3 illustrates the final stage, including *CR*'s 56 through 59. Here the records are available for six animals only, Animal *Y* having become emotionally unsuitable for continued experimentation. The principal feature of this figure is an increase in the duration of each flexion; the increase, however, is not due to an immobile suspension of the foreleg in mid-air but represents, according to the protocols, an oscillating vertical movement.

Figure 4 shows tracings from the polygraph records of the remaining two animals. These are typical of the oscillation which occurs during the final or adaptive stage.

The results are in harmony with the principle of contiguity if the progress from one stage to another is interpreted in terms of enlistment of additional cues. One of these cues is the tactile and kinesthetic stimulation which takes place when the animal's toes come in contact with the ground. Normally it evokes the positive supporting reaction; and, during the initial stage, such stimulation occurs maximally when the limb is replaced after the first brief *CR*. Since, however, it is invariably followed by the unconditioned response, its function in the maintenance of posture is replaced by its new rôle as a signal for the contraction of flexors. The effect is demonstrated by the successive flexions observed during the intermediate stage.

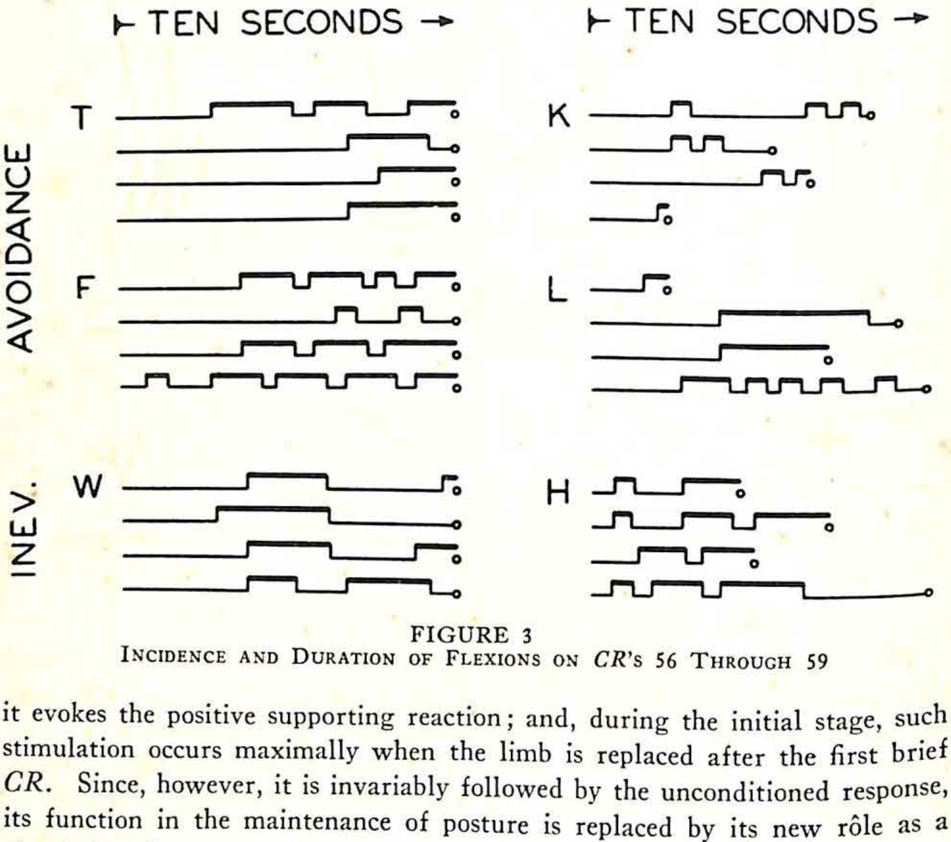

FIGURE 3
INCIDENCE AND DURATION OF FLEXIONS ON *CR*'S 56 THROUGH 59

Development of the final stage may be attributed to enlistment of still another cue: the kinesthetic return from contraction of the extensors in the downward movement of the limb. In a conditioning experiment such feedback regularly precedes the tactile-kinesthetic stimulation resulting from actual contact. Assuming that the latter has now become a cue for flexion, the former may be regarded as a conditioned stimulus of second order.

On the basis of this interpretation one might expect that avoidance would be achieved by an up-and-down movement sequence, which, involving the function of antagonists, is facilitated by successive induction (4). The reactions observed in the final stage of conditioning appear to validate this inference.

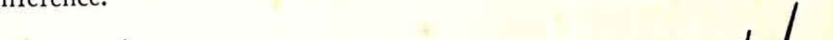

FIGURE 4
POLYGRAPH RECORDS OF FLEXIONS WHICH OCCUR DURING THE FINAL STAGE OF CONDITIONING

## D. DISCUSSION

The results suggest that refutation of the principle of contiguity, derived from phenomenal discrepancies between *CR* in its final form and *UR*, may be premature. On the other hand, formulations of the principle of contiguity, which disregard or fail to specify the nature of intervening events, are limited in predictive capacity and vulnerable to legitimate criticism.

## E. SUMMARY

Observation of nine goats in a conditioning experiment shows that the *CR* reaches its final form through a stable and uniform sequence of changes. The nature of these changes suggests that stimuli emanating from the animal's own movements are enlisted as part of the *CS*-complex. The results are interpreted as being in harmony with the principle of contiguity.

## REFERENCES

1. CULLER, E., FINCH, G., GIRDEN, E., & BROGDEN, W. J. Measurement of acuity by the conditioned-response technique. *J. Gen. Psychol.*, 1935, **12**, 223-227.
2. HILGARD, E. R., & MARQUIS, D. G. Conditioning and Learning. New York: Appleton-Century, 1940.
3. HUMPHREY, G. Thinking. New York: Wiley, 1951.
4. SHERRINGTON, C. S. The Integrative Action of the Nervous System. New Haven: Yale Univ. Press, 1947.
5. SKINNER, B. F. The Behavior of Organisms. New York: Appleton-Century, 1938.
6. TOLMAN, E. C. Purposive Behavior in Animals and Men. New York: Century, 1932.

*Department of Psychology*
*Cornell University*
*Ithaca, New York*

*The Journal of Genetic Psychology*, 1957, **90**, 89-95.

# TERMINAL EXTINCTION IN A MULTIPLE-*T* MAZE WITHIN A HETEROGENEOUS ENVIRONMENT*

*Department of Psychology, University of Texas*

W. LYNN BROWN[1] AND EDWARD S. HALAS

## A. LITERATURE AND PROBLEM

In contemporary learning theory the principle of secondary reinforcement is frequently used to explain the resistance of a learned response to extinction. Hull (3) in 1943 assumes that "*a receptor impulse will acquire the power of acting as a reinforcing agent if it occurs consistently and repeatedly within 20 seconds or so of a functionally potent reinforcing state of affairs, regardless of whether the latter is primary or secondary*" (pp. 94-95). The position assumes that extinction will be retarded in the presence of stimuli which have been repeatedly associated with reinforcement.

Seward and Levy (6) have introduced the concept "latent extinction" in a study of non-response extinction. They assume that sign-learning theory can readily account for non-response extinction while reinforcement theory will have difficulty in giving a satisfactory explanation. In an experiment similar to that of Seward and Levy, Deese (1) found positive results for non-response extinction, i.e., that animals which were placed in the empty goal-box before extinction training, extinguished more rapidly than did the animals without this procedure. These results stand in opposition to those of Scharlock (5) who obtained negative results in a study of non-response extinction. More recently Moltz (4) concludes that latent extinction is brought about by a reduction of secondary reinforcing properties of a relevant stimulus (food cup). This position supports a possible reinforcement interpretation.

Even though Hull postulated non-response extinction and Moltz does not mention the possibility of response extinction, it is difficult to attribute the positive results that Seward and Levy (3) and Deese (1) got to non-response extinction *per se*. It can be readily observed that *S*s make many responses in the terminal segment of a maze and in the goal-box, some are involved in the consumption of food, but many are not. The *S* may stop be-

---

*Received in the Editorial Office on November 14, 1955.

[1] This investigation was supported in part by a grant to the senior author by the Research Institute of the University of Texas.

fore entering the goal-box and examine it; he may explore the interior by sniffing, looking, and climbing. It is quite conceivable that some of these responses may become associated with reward during the training series. Also available to *S* during this period are many stimuli in the area of the final segment of the maze and goal-box plus extra-maze stimuli from the room and kinesthetic stimuli from responses. Therefore, it is our contention that non-response extinction may actually be extinction of terminal responses that have been previously rewarded or closely associated with reward. It is also our contention that the more explicit these terminal responses are and the more stimuli available to elicit these terminal responses, the greater will be the effect of extinction from a given number of extinction trials.

Contrary to the opinion of Seward and Levy, reinforcement theory can account for non-response extinction. In fact, Hull (2) predicted and postulated non-response extinction when he wrote:

> When an excitatory tendency in a behavior sequence encounters a situation which makes the execution of the act impossible (Definition 3), the excitatory tendency in question undergoes a diminution in strength with a limit below the reaction threshold (Definition 2), this diminution extending in considerable part to other excitatory tendencies which may be operative at the same time or for some time thereafter (p. 17).

The present experiment was designed to study the effect of a removal of a visual stimulus from the terminal area while leaving intact other stimuli plus responses formerly associated with the terminal area (other than eating).

## B. MATERIALS AND METHODS

### 1. *Apparatus*

A 14-unit multiple-*T* maze of the Tolman-Honzik design was used in the experiment. The units were 24 in. x 2 in. and stood 15 in. from the floor. The first 13 units were painted mid-gray while the 14th and the goal-box were painted white. The goal-box was 9 in. x 9 in. x 6 in. The entrance was 2 in. wide and extended from the floor of the box to the top. The box was not covered. To avoid the possibility of tracking by smell, the maze units were washed daily with hot water.

### 2. *Subjects*

Eighteen experimentally naïve, male rats of Wistar strain, ranging in age from 100 to 130 days, were used.

### 3. *Experimental Room*

A room, 13 ft. 4 in. x 16 ft. 9 in. with a 9-ft. ceiling was used. One window is centered on the west wall and one on the east wall. Two doors are on the south and one on the north.

### 4. *Preliminary Training*

*S*s were placed on a 24-hr. feeding regimen for eight days during which time they were given a total of 20 trials on a straight pathway to a natural wood goal-box with wet mash inside. This training was in the colony room.

### 5. *Maze Training*

*S*s were given one run a day for nine days. Errors and time were both recorded. An error was recorded when an *S* entered a cul with four feet or his nose reached the end of the cul. Culs were about 6 in. deep. Retracing was not permitted and only one error per cul was recorded. Time was taken from the placing of *S* on the first unit until all feet were in the goal-box.

When an *S* entered the goal-box, the 14th unit was moved out of position to prevent retracing. *S*s were left in the goal-box for one minute after each run and were allowed to eat a wet mash, after which they were returned to their home cage and were fed one large pellet each of Purina Laboratory Chow.

### 6. *Results of Maze Training*

The learning curves for time and error are shown in Figures 1 and 2. A total of three errors were made by the ninth day by the 18 *S*s. The average running time was 17.11 seconds.

### 7. *Extinction Procedure*

On the 10th day the *S*s were divided into two matched groups of nine each on the basis of time and error score. Both groups were given eight extinction trials. The extinction series for Group 1 was on the white 14th unit and white goal-box used during the training series. Group 2 ran on a black unit with a black goal-box of the same dimensions as the white one used by Group 1. Extinction procedure was as follows: The maze was set up in the position it occupied during training with the exception of the 13th unit which was pulled out of position to prevent retracing during extinction trials. Each *S* was placed at the end of the 14th unit and allowed to run the unit and to enter the goal-box where it was detained for one minute without food. Eight extinction trials were given in succession with 15 seconds separating the trials. One minute after the eighth trial the *S* was placed at the start of the training maze and allowed to run the entire maze. Time and error scores were recorded during this test run.

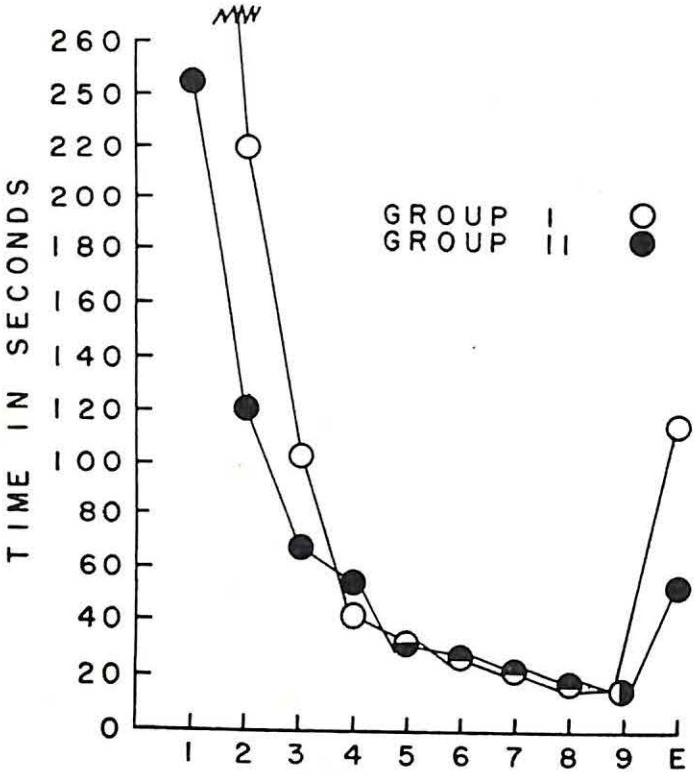

FIGURE 1
RUNNING TIME DURING LEARNING AND EXTINCTION

## C. RESULTS AND DISCUSSION

Even though the two groups began their extinction series with the same average time and nearly the same error score (see Figures 1 and 2) they separate very decidedly on both measures on the test run following the extinction trials. Group 1 had an average running time of 116.33 seconds on the test run and Group 2 54.22 seconds. This difference between the groups is significant at the 5 per cent level ($t = 2.37$).

Group 1 made one error on the final day of training and 43 following the extinction trials. This difference is significant well beyond the 1 per cent level ($t = 5.35$). Group 2 made two errors on the final day of training and 29 following extinction, which is also significant beyond the 1 per cent level ($t = 3.88$). A $t$ of 1.44 was obtained between the groups on the test run.

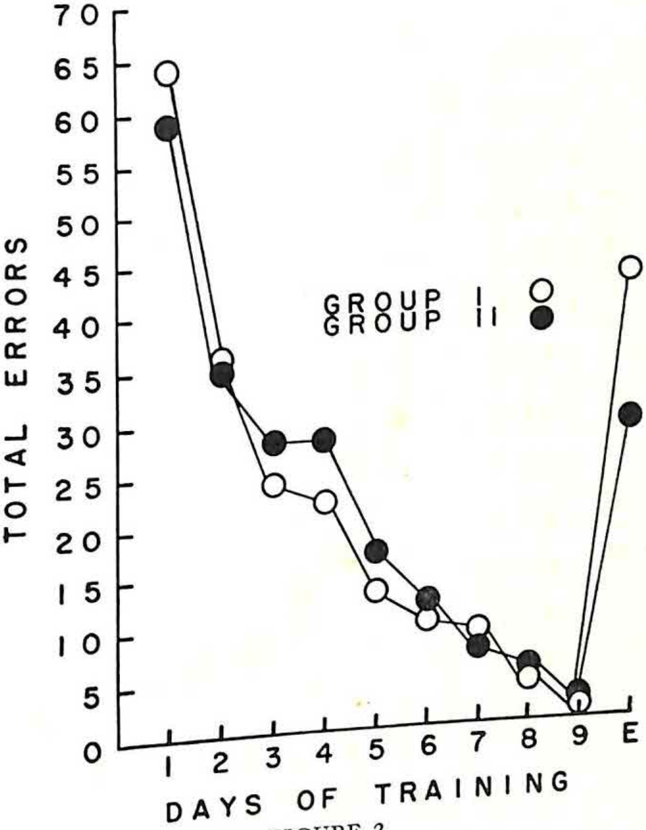

FIGURE 2
CURVES FOR LEARNING AND EXTINCTION

Both groups had a running time average of 17.11 seconds on the ninth day of training. On the test run, Group 1 differed significantly from this figure beyond the 1 per cent level ($t = 3.78$) and Group 2 likewise by a $t$ of 6.46.

The results from the test run show that the groups separated on both time and error measures. The extinction trials for Group 1 were on the final unit and goal-box used in the training series, hence secondary reinforcement was the same in this area as during training. The white pathway and white goal-box, room position of the final unit and goal-box, windows, doors and walls were the same. In addition to the visual stimuli, both intra- and extra-maze, available to the *S*s, there were also kinesthetic and contact stimuli from running on the pathway plus stimuli derived from activity in the goal-box—primary reinforcement or no primary reinforcement. It may be assumed, furthermore, that *S* receives stimuli from handling by the experimenter which become a part of the complex of stimuli associated with the problem. This was not the case, however, for Group 2 which ran a black pathway and entered a black goal-box. The complex of stimuli, other than intra-maze visual stimuli, formerly associated with primary reinforcement, were available to the *S* if it could use them. The results clearly show a deterioration of performance of both groups following extinction trials, with Group 1, the group with more secondary reinforcement during extinction trials, showing greater deterioration. Since the black pathway and goal-box used in the extinction series of Group 2 had not been associated with primary reinforcement, hence possessing no visual secondary reinforcing values, one must assume that other stimuli were involved in extinction.

Extinction on the black unit and goal-box was less effective than on the white unit and goal-box, i.e., the *S*s in Group 2 made fewer errors and ran faster than those in Group 1. It must be recalled here that extra- and intra-maze cues, other than color, were intact for Group 2. Apparently in the present experiment a reduction in secondary reinforcing cues strengthened resistance to extinction.

Many of the *S*s in Group 1 retraced in Section 13 when they reached the 14th unit and also frequently made an error in Cul 14 upon entering the terminal section. This type of behavior was not observed in Group 2. They entered the 14th unit without hesitation and made no errors after entering.

## D. SUMMARY

Rats were trained to run a multiple-*T* maze with the first 13 units painted a mid-gray and the 14th and goal-box painted white. Group 1 was given eight extinction trials on the white section and then tested on the original maze, and Group 2 was given the same number of extinction trials on a black terminal section and also tested on the original maze. Both groups showed a deterioration. The results were discussed in terms of secondary reinforcement.

## REFERENCES

1. DEESE, J. The extinction of a discrimination without performance of the choice response. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 362-366.
2. HULL, C. L. Mind, mechanism and adaptive behavior. *Psychol. Rev.*, 1937, **44**, 1-32.
3. ———. Principles of Behavior. New York: D. Appleton-Century, 1943.
4. MOLTZ, H. Latent extinction and the reduction of secondary reward value. *J. Exper. Psychol.*, 1955, **49**, 395-400.
5. SCHARLOCK, D. P. The effects of a pre-extinction procedure on the extinction of place and response performance in a *T*-maze. *J. Exper. Psychol.*, 1954, **48**, 31-36.
6. SEWARD, J. P., & LEVY, N. Sign learning as a factor in extinction. *J. Exper. Psychol.*, 1949, **39**, 660-668.

*Department of Psychology*
*University of Texas*
*Austin 12, Texas*

*The Journal of Genetic Psychology*, 1957, **90**, 97-101.

# TERMINAL EXTINCTION IN A MULTIPLE-*T* MAZE WITHIN A HOMOGENEOUS ENVIRONMENT*

*Department of Psychology, University of Texas*

W. LYNN BROWN[1] AND EDWARD S. HALAS

## A. PROBLEM

In an investigation on terminal extinction in a heterogeneous environment, reported in this Journal by Brown and Halas (1), *S*s were trained on a 14-unit multiple-*T* maze with the first 13 units painted mid-gray and the final unit and goal-box painted white. The *S*s in Group 1 were given eight extinction trials on the white unit and white goal-box while the *S*s in Group 2 were given extinction trials on a black unit with a black goal-box. When the *S*s of both groups were run on the training maze following extinction, the investigators found a marked deterioration in the performance of both groups in terms of time and error, with Group 1, the group with more secondary reinforcement during extinction trials, showing greater deterioration. Brown and Halas discussed the results in terms of secondary reinforcement.

The present experiment was designed to answer, if possible, some of the questions left unanswered in the first experiment. Since Groups 1 and 2 performed differently from each other following extinction training and also significantly differently from their pre-extinction training, it is assumed that a further reduction in specific secondary reinforcing cues might narrow the differences following extinction training between the groups and also reduce deterioration following extinction training.

## B. MATERIALS AND METHODS

### 1. *Experimental Room*

The second study was conducted in the same manner as was the first, with the exception of the experimental room, which was a homogeneous dome 18 ft. in diameter with 5½-ft. side walls arched to an 8-ft. center. It is constructed symmetrically of wooden ribs and covered with a fine mesh screen painted silver gray, hence gives a homogeneous surface radically symmetrical about the geometric center. A 150-w. bulb was used in the center

---

*Received in the Editorial Office on November 14, 1955.

[1] This investigation was supported in part by a grant to the senior author by the Research Institute of the University of Texas.

of the dome with a 12-in. disk placed 4 in. below the bulb to reflect the light to the ceiling of the dome. This produced indirect lighting during the experiment.

### 2. *Subjects*

Eighteen experimentally naïve female rats, ranging in age from 120 to 150 days old, were used.

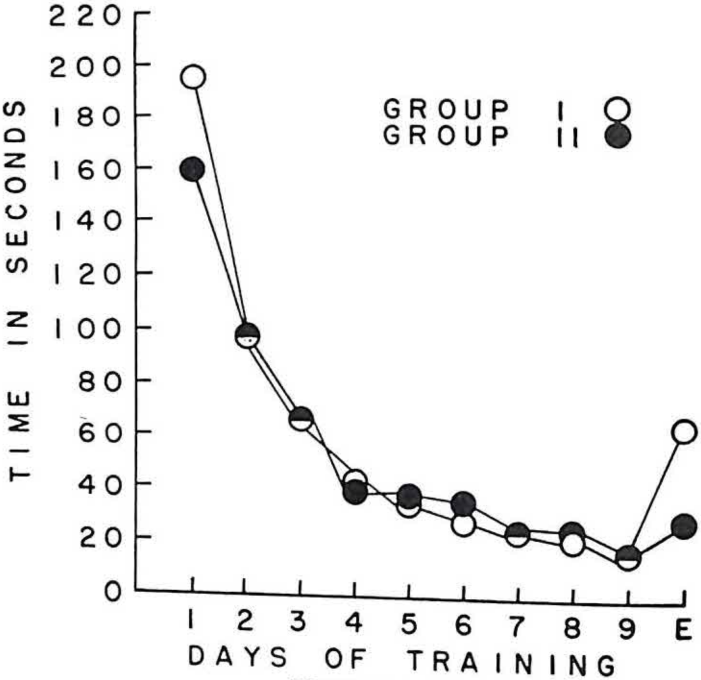

FIGURE 1
RUNNING TIME DURING LEARNING AND EXTINCTION

### 3. *Training*

Both preliminary and experimental training were the same as in the first study.

### 4. *Results of Maze Training*

A total of seven errors was made on the last day of training by the 18 *S*s. The average running time was 17.22 seconds. Learning curves for time and errors are shown in Figures 1 and 2.

### 5. *Extinction Procedure*

Before the eight extinction trials were given on the 10th day the *S*s were divided into two matched groups of nine each on the basis of time and error score. Group 1 ran on a white pathway to a white goal-box and Group 2 on a black pathway to a black goal-box. The only difference in extinction procedure in the present study and the former was in the time interval on the extinction pathway. If an *S* did not enter the goal-box in the present study after one minute on the pathway, it was picked up and placed in the box, whereas in the former study a time limitation was not used.

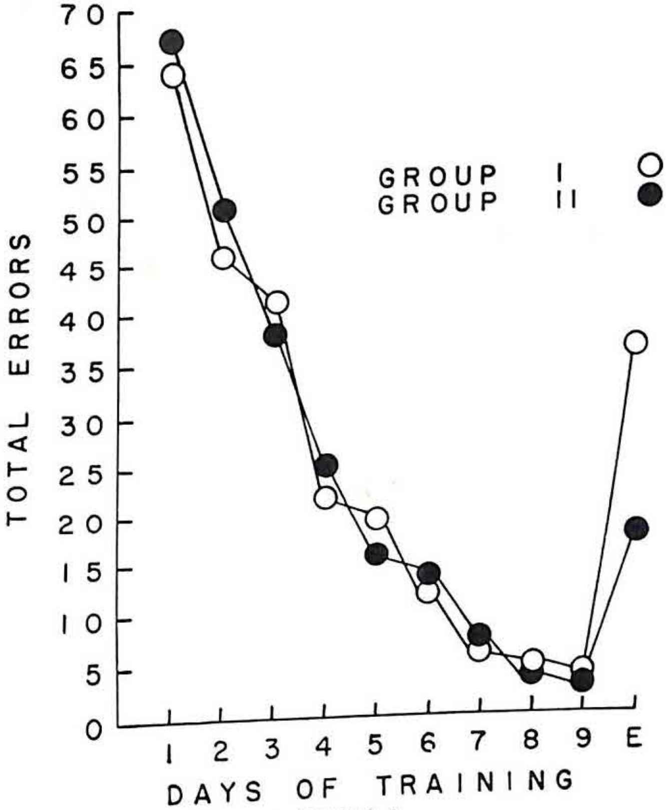

FIGURE 2
Curves for Learning and Extinction

## C. Results and Discussion

The performance of *S*s on the training maze following the extinction series is presented in Figures 1 and 2 for both time and error. The figures show that prior to extinction the groups were nearly the same in performance. Both groups had a final running time average of 17.22 sec. On the test trial, however, the groups differed significantly from their final running time average and also from each other. Group 1 averaged 61.22 sec. on the test which is significant beyond the 1 per cent level ($t = 4.84$). Group 2 averaged 29.22 on the test which is also significant beyond the 1 per cent level ($t = 4.36$). The difference between the groups was significant beyond the 1 per cent level ($t = 4.02$).

When the groups are considered for errors on the test trial, again we find that they differ in performance. Group 1 made 4 errors on the final day of training and 36 on the test run. This difference is significant beyond the 1 per cent level ($t = 5.07$). Group 2 made 3 errors on the final run and 18 on the test; a difference also significant at the 1 per cent level ($t = 3.40$). The difference between the groups was at the 7 per cent level ($t = 2.50$).

Again we have observed a deterioration of performance on a test trial following eight extinction trials. A comparison of the results of the first study (1, Figures 1 and 2) with the results of the present study (Figures 1 and 2) will show that the magnitude of deterioration of performance following extinction was greater in both time and error for the *S*s in the first study. The *S*s in Group 1 in the first study in a heterogeneous environment made 43 errors and had a running time average of 116.33 seconds following extinction while the *S*s in the present study made 36 errors and had a running time average of 67.22 seconds. In a comparison of Group 2 in the first study with Group 2 in the second we find that Group 2 (first study) made 29 errors and had a running time average of 54.22 seconds while the *S*s in Group 2 (second study) made 18 errors and had a running time average of 29.22 following eight extinction trials. Even though the present experiment was conducted in a homogeneous experimental dome which greatly reduced or eliminated specific extra-maze cues which might serve as secondary reinforcing stimuli, the *S*s still differed significantly beyond the 1 per cent level on the test run in both time and error from their final training performance.

When the intra-maze cues were reduced in the first experiment for Group 2 on the extinction trials, we found less deterioration in performance than for Group 1 which ran the extinction trials on the training maze segment. With a further reduction of cues through the use of an experimental dome in the present experiment we observed a reduction in both error and average running time from the results obtained in the first experiment. Apparently a reduction of cues in the present experimental design, intra- or etxra-maze, is followed by increased resistance to extinction.

## D. Summary and Conclusion

The general trend in deterioration of performance following the extinction trials, in both time and error measures, was the same for both groups, with Group 1 showing greater deterioration. Since the only difference in the complex of stimuli presented to the *S*s in the two groups was in the intra-maze visual stimuli, one must assume that this factor accounted for the differences in performance. One must assume, furthermore, that the difference in performance between the *S*s in the two experiments is attributable to differences in extra-maze stimuli, the first study being conducted in a heterogeneous environment and the second in a homogeneous. The experimenters also assume that if a reduction in responses in the extinction area was introduced, thus reducing kinesthetic stimuli closely associated with reinforcement, an additional reduction in time and error would result in both experimental situations with the homogeneous showing a greater reduction. This assumption is under investigation.

## Reference

1. Brown, W. L., & Halas, E. S. Terminal extinction in a multiple-*T* maze within a heterogeneous environment. *J. Genet. Psychol.*, 1957, **90**, 89-95.

*Department of Psychology*
*University of Texas*
*Austin 12, Texas*

*The Journal of Genetic Psychology*, 1957, **90**, 103-107.

# RELATIONSHIP BETWEEN EMOTIONAL ELIMINATION AND THE PERSISTENCE OF AN ANXIETY MOTIVATED RESPONSE*

*Department of Psychology, Carnegie Institute of Technology*

---

K. E. Moyer[1]

---

## A. Problem

Emotional elimination has been considered to be a reasonably accurate indication of emotionality in the rat. Hall has suggested that it is related to fear (4), and Billingslea (1) has indicated that the more emotional rat (as determined by the Hall open field test) is more timid in the home cage and in the presence of the experimenter. Tryon, Tryon, and Kuznets (10) developed a rating scale designed to quantify emotional behavior in the rat. They report that emotional elimination in the female as determined by Hall's test correlates highly (.83) with their rating scale. In the male, however, the correlation reported is only .34. These authors conclude that their data generally support Hall's contention that emotional defecation is related to emotional excitability.

Whether or not high elimination in an open field situation is a valid indication of emotional excitability was initially questioned by O'Kelley (9). More recently Bindra and Thompson (2) found that emotional elimination was not correlated with either fearfulness or timidity as they measured it. The results of Hunt and Otis (5) agree that emotional defecation as measured by the open field test is not correlated with timidity. One of their experiments shows, however, that "defecation can serve to indicate conditioned emotional disturbance of the 'fear' or 'anxiety' type." It was observed to be a consistent component of a conditioned emotional response. They suggest that defecation may be clearly related to emotional behavior in the rat, but the relationship becomes reliable "only when the emotional disturbance is substantial and reasonably intense."

This experiment is an attempt to study further the relationship between the open field test of emotionality and another test of the rat's susceptibility to emotional disturbance.

---

*Received in the Editorial Office on December 5, 1955.

[1] The expense involved in this study was, in part, met by a grant from the Humanistic and Social Science Research Fund of the Carnegie Institute of Technology.

It has been suggested by Mowrer (6, 7) that persistence of an avoidance response to a *CS* which has been previously associated with shock is due to the fact that the *CS* elicits in the animal an uncomfortable autonomic response (anxiety) as well as a running or avoidant response (see also 3 and 8). If this is a valid argument, it would be suspected that individual differences in emotionality would be correlated with individual differences in susceptibilty to the formation of strong conditioned anxiety responses.

## B. METHOD

### 1. *Subjects*

In this study, 18 naïve female hooded rats were used. They were 90 to 110 days old at the beginning of experimentation.

### 2. *Apparatus*

The modified open field consisted of a porcelain table top 27 by 40 inches, with sides of composition board 15 inches high. A 200-watt bulb was placed 23 inches above the table and the table top was 32 inches above the floor.

The apparatus used to establish the conditioned emotional response has been described in detail elsewhere (8). Briefly, it consisted of an alley 10 inches wide, 35 inches long, with sides 16 inches high. There was a goal box in the front left-hand corner. The floor of the main alley was composed of two sections of brass rods, each section of which could be electrified separately, or both could be charged at the same time. The goal box was provided with a guillotine door. Shock was provided by an electronic shocker with sufficient resistance built into the circuit to minimize individual resistance differences in the *S*s.

### 3. *Procedure*

All *S*s were treated in the same way during the open field test and the training trials. Food and water was available in the home cages at all times. Each animal was placed in the center of the open field for a four-minute period once a day for four days. A record was kept of the number of times the *S* urinated and the number of fecal bolli produced. Between trials the sides were removed and the open field was washed with warm water and then allowed to dry.

Each *S* was given 10 training trials in the conditioning apparatus. A training trial consisted of placing the *S* in the right-hand side of the box about three inches from the right and facing that end. The goal box door was closed. After 10 seconds, shock was applied to the entire grid floor and the goal box door was opened. Shock remained on until the *S* achieved the goal box and thus escaped it. Approximately 10 seconds elapsed between trials while the response was being recorded by *E*.

The animals were divided into two groups for the extinction trials because previous research has shown that these two extinction procedures produce different results. It was felt that these procedures might correlate differently with Hall's test of emotionality.

The extinction trials for Group 1 were the same as the training trials except that the grid was never charged, and the goal box door was always open. It was assumed that the response of running to the goal box was extinguished when the S remained outside it for two minutes.

During the extinction trials for Group 2 the left half of the grid was always charged and the right half never was. Thus, it was necessary for the *S* to run from a safe area through a shock area to get to the goal. The same criterion of extinction was used for this group, and the goal box door was always open during the extinction trials.

Each animal was given 30 extinction trials immediately after the training trials. If the *S* did not meet the criterion within 30 trials, it was given 40 trials a day on subsequent days until it did. Between extinction trials 10 seconds elapsed.

## C. RESULTS

There were three measures of emotional elimination used in this study: (*a*) number of days on which a *S* defecated, (*b*) number of days on which a *S* either defecated or urinated, and (*c*) a combined defecation and urination score. The last measure was computed as follows: the total number of times that the *S* urinated was added to the number of fecal bolli produced in the open field situation during all four days.

A rank order correlation was computed for each group between the number of trials to extinction and each of the measures of emotional elimination. These correlations are presented in Table 1. None of them is significant.

## D. DISCUSSION

The open field test is presumed to be a stimulus which, because of its strangeness, inherently evokes an autonomic response in the rat which is measured by the amount of elimination produced in the situation. The more emotional rat, by definition, is the one which responds autonomically to the stimulus for the longest time (greatest number of days). In this experiment, according to the Mowrer hypothesis, the neutral cues of the box because they were associated with shock acquired the power to elicit an autonomic or anxiety response in the rats. In other words, the *S*s learned two responses: first, the response of running to the goal when placed in the box, and second, an autonomic response referred to as anxiety. Again, according to the Mowrer hypothesis, the autonomic response is reduced when the *S* achieves the goal thus escaping the cues which cause it. Since the anxiety state is uncomfortable for the animal, escape from it constitutes a reward. Inasmuch as the emotional animal, as defined above, responds longer to an anxiety producing stimulus, it would be expected that the more emotional rat would take longer to extinguish the response of running to the goal. Hence we would expect a positive relationship between emotionality and the number of trials to extinction of an anxiety motivated response. Such a relationship was not found in this study.

TABLE 1

RANK ORDER CORRELATIONS BETWEEN NUMBER OF TRIALS TO EXTINCTION FOR THE SHOCK AND NON-SHOCK GROUPS AND THREE MEASURES OF EMOTIONAL ELIMINATION

| | Shock group number of trials to extinction | Non-shock group number of trials to extinction |
|---|---|---|
| Number of days defecating | —.12 | .33 |
| Number of days defecating or urinating | .14 | .15 |
| Combined defecation and urination scores | .24 | —.19 |

Since, in this situation, the anxiety response is a learned one, it may be that different levels of learning occurred in different animals in the same number of trials. Thus, different levels of anxiety may have been generated by the cues of the box irrespective of the inherent emotionality of the *S*. Consequently a more emotional animal might be responding to a less intense autonomic stimulation and would extinguish in fewer trials than a less emotional *S* which was responding to autonomic stimulation of greater intensity. This may be one explanation of the lack of correlation found in this study.

Considering the findings of Bindra and Thompson (2) and Hunt and Otis (5), the possibility must also be entertained that the lack of correlation is due to the fact that the open field test is not a valid measure of emotionality in the rat.

## E. SUMMARY

Emotionality scores were obtained for two groups of nine rats in the open field test. Both groups were then taught to escape shock by running to a goal box. The extinction trials for Group 1 were the same as the training trials except that shock was never administered. For Group 2 it was necessary for the animal to run from a safe area on the grid through a shock area to get to the goal.

The emotionality scores were correlated with the extinction scores. None of the correlations were significant.

## REFERENCES

1. BILLINGSLEA, F. Y. The relationship between emotionality and various other salients of behavior in the rat. *J. Comp. Psychol.,* 1941, **31**, 69-77.
2. BINDRA, D., & THOMPSON, W. R. An evaluation of defecation and urination as measures of fearfulness. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 43-45.
3. FARBER, I. E. Response fixation under anxiety and non-anxiety conditions. *J. Exp. Psychol.*, 1948, **38**, 111-131.
4. HALL, C. S. Emotional behavior in the rat: I. Defecation and urination as measures of individual differences in emotionality. *J. Comp. & Physiol. Psychol.*, 1934, **18**, 385-403.
5. HUNT, H. F., & OTIS, L. S. Conditioned and unconditioned emotional defecation in the rat. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 378-382.
6. MOWRER, O. H. Anxiety reduction and learning. *J. Exp. Psychol.*, 1940, **27**, 497-516.
7. ———. Learning theory and the neurotic paradox. *Amer. J. Orthopsychiat.*, 1948, **18**, 571-610.
8. MOYER, K. E. A study of some of the variables of which fixation is a function. *J. Genet. Psychol.*, 1955, **86**, 3-31.
9. O'KELLEY, L. I. The validity of defecation as a measure of emotionality in the rat. *J. Gen. Psychol.*, 1940, **23**, 75-87.
10. TRYON, R. C., TRYON, C. M., & KUZNETS, G. Studies in individual differences in maze ability: IX. Ratings of hiding, avoidance, escape, and vocalization responses. *J. Comp. & Physiol. Psychol.*, 1941, **32**, 407-435.

*Department of Psychology*
*Carnegie Institute of Technology*
*Pittsburgh, Pennsylvania*

*The Journal of Genetic Psychology*, 1957, **90**, 109-117.

# AUDIOGENIC SEIZURES: INSULIN AND HIGH BLOOD SUGAR*

*Cornell University, General Electric Advanced Electronics Center*

Alfred Steinschneider[1]

## A. Problem

Since the report by Maier (10), a great deal of research has been conducted in an attempt to elucidate those parameters affecting the incidence and latency of audiogenic seizures. The behavioral sequence noted during an audiogenic seizure and first reported by Donaldson (1) can be initiated by the presentation of an intense, high frequency sound.

A short time after the onset of the auditory stimulus, the seizure-prone animal runs vigorously and apparently without direction (running phase). This is often followed by a tonic, a clonic, or a tonic-clonic phase. The sequence is concluded by a comatose phase, which is characterized by extreme passivity. The seizure proves fatal to some animals; however, the majority gradually recover.

The hypothesis was formulated that one of the mechanisms resulting in the audiogenic seizure is essentially the same as that for insulin-produced seizures. This hypothesis suggests that, during the presentation of the auditory stimulus, the seizure-prone animal secretes an "excessive" amount of insulin into his blood stream and that this hormone contributes to the production of the seizure.

The initial experiment consisted of determining the seizure susceptibility of seizure-prone mice following the destruction of the beta cells within the pancreas. It was reasoned that if susceptibility to audiogenic seizures is a function of the amount of insulin secreted during the presentation of the auditory stimulus, then decreasing this amount would decrease seizure susceptibility. However, since a decrease in the amount of insulin secreted is accompanied by an increase in the blood sugar level, a second experiment was conducted. In this latter experiment, the blood sugar level was increased without disturbing the insulin producing capacity of the pancreas.

---

*Received in the Editorial Office on December 8, 1955.

[1] The writer wishes to express his appreciation to Dr. Morley Kare of the New York State Veterinary College for his guidance throughout this study, and for the privilege of utilizing his laboratory facilities.

## B. Method

### 1. *Apparatus and General Testing Procedure*

The seizure-producing auditory stimulus was at 9,000 cps and 94 ± 3db (reference level $10^{-16}$ watt square centimeter) average sound pressure. This tone, produced by a Hewlett-Packard Audio-oscillator Model 200 CD and amplified by a Heathkit amplifier Model A-6A, was fed into a speaker housed in a semi-sound-proof box (see Figure 1). An albino mouse was placed in a tubular wire-mesh cage 3½ inches long by 2¾ inches in diameter which, in turn, was suspended on two wooden dowels within the box. The center of the cage was four inches from the speaker. Observation of the mouse was made through a glass window in the front of the box. The inside of the box was illuminated by a 7½-watt milk-white bulb hung from the ceiling.

Each exposure period lasted a maximum of 60 seconds. If the animal went into seizure during this period, the stimulus was terminated at the onset of the clonic phase. Animals were tested individually.

### 2. *Experiment I*

The strain of albino mice used in both experiments of this study was obtained from an original mating pair received from Pennsylvania State College.[2] This particular strain of mice was chosen because of its high seizure susceptibility and uniformity of seizure pattern (4, 5). The seizure pattern of these animals can be schematized as follows: running phase, clonic phase, comatose phase.

A total of 48 mice started Experiment I—23 males and 25 females. These animals were between the ages of one and two months. Throughout the experiment, the animals were offered food and water *ad libitum*.

Alloxan monohydrate was used to destroy the beta cells of the pancreas (8). An effective dosage was found to be 180 mgm/kgm body weight and, consequently, a 3.6 per cent water solution of alloxan was introduced intravenously into the animal through one of the tail veins.

It was intended that each mouse be given three exposures to the auditory stimulus. The first exposure served to eliminate the few animals that were not seizure susceptible. The remaining two exposures occurred following the placement of the seizure-susceptible animals into one of two groups.

On the day following the initial seizure test, one group of 25 animals received alloxan (Alloxan Group), while a second group ($N = 23$) was given an equivalent volume of physiological saline (Saline Group).

---

[2] These animals were obtained through the kind coöperation of Mable and Hubert Frings, Department of Zoölogy and Entomology, the Pennsylvania State College.

To determine the effectiveness of alloxan in destroying the insulin producing cells, the presence of sugar in the urine was determined for each animal

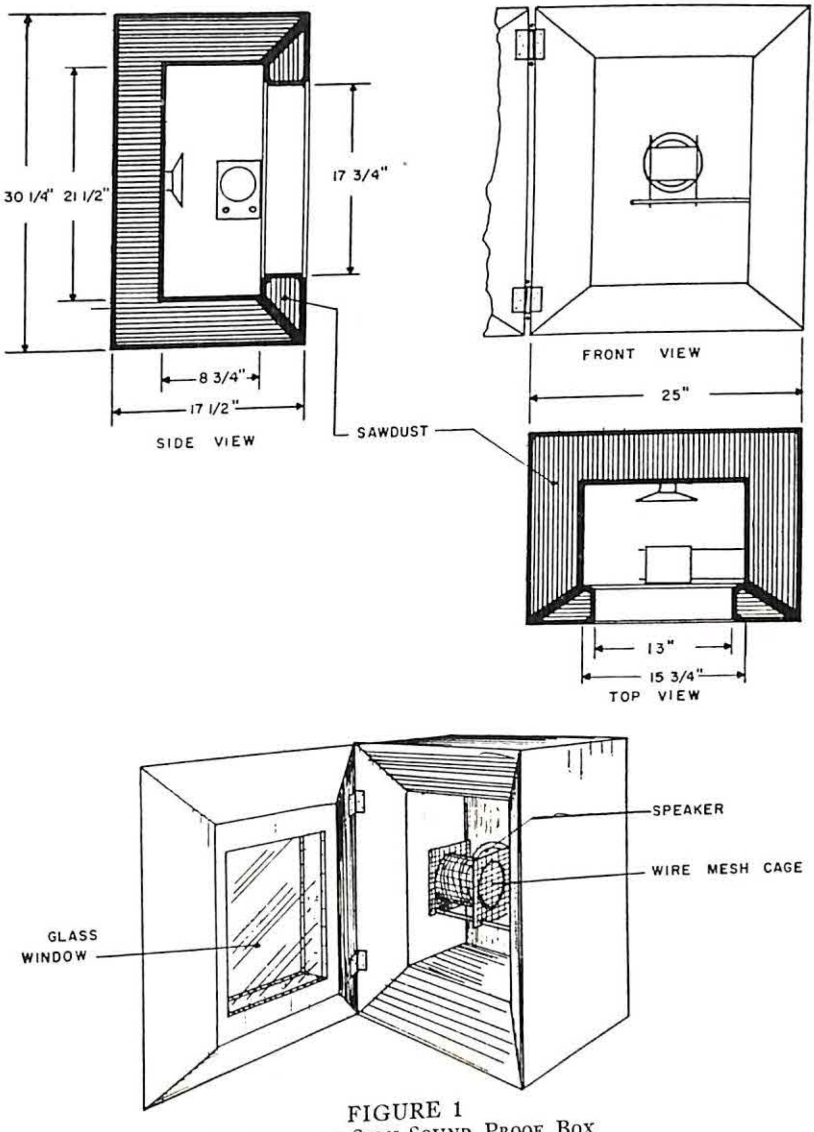

FIGURE 1
Diagram of Semi-Sound Proof Box

on the two days prior to the first seizure session, two days following the administration of alloxan or physiological saline, and on the day prior to the third seizure session. Only those animals who demonstrated a marked glucosuria were retained in the Alloxan Group. Two mice were discarded for failure to exhibit glucosuria two days after the administration of alloxan. Clinitest tablets were used to detect the presence of sugar in the urine.

In addition to the two mice discarded, five other animals died within two days following administration of alloxan. Thus, 18 animals remained in the Alloxan Group for the first post-alloxan seizure test. Prior to the second post-alloxan seizure test, four more animals died, leaving 14 mice for the last seizure exposure.

The daily sequence of sessions can be summarized as follows:

Days 1 and 2—determination of urine sugar.
Day 3 —pretreatment seizure test.
Day 4 —injection of alloxan monohydrate or physiological saline.
Days 5 and 6—determination of urine sugar.
Day 7 —first post-treatment seizure test.
Days 8 and 9—rest.
Day 10 —determination of urine sugar.
Day 11 —second post-treatment seizure test.

### 3. *Experiment II*

A total of 14 naïve mice—nine males and five females—were selected to take part in this experiment. Only those animals who demonstrated the complete seizure pattern during their first exposure to the auditory stimulus were chosen. These animals were approximately 30 days old at the time of the first stimulus exposure.

Subsequent to the first seizure exposure, each animal was tested on two more occasions. Testing sessions were spaced three days apart.

Three minutes prior to the second test exposure, one-half of the animals received approximately 0.1 cc of a 5 per cent dextrose solution. The remaining animals were given an equivalent volume of physiological saline. These solutions were introduced intravenously into one of the tail veins. This same procedure was followed on the third test exposure, except that those animals who received the dextrose solution on the second exposure now received the saline solution, whereas those previously administered the saline solution were given the dextrose solution.

### 4. *Measures of Seizure Susceptibility*

Two measures of seizure susceptibility have been used when dealing with audiogenic seizures (3); frequency of seizures within a group and latency. In this experiment, frequency and latency were determined for both the running phase and the clonic phase. The latter measure, defined as the time interval between the onset of the auditory stimulus and the onset of the phase in question, was obtained by means of two electric stop watches (one for each phase).

A decrease in seizure susceptibility would be reflected in both a decrease in the frequency of seizures and an increase in seizure latency. The reverse would indicate an increase in seizure susceptibility.

## C. RESULTS

### 1. *Experiment I*

Table 1 contains the results obtained on the first post-treatment seizure test. Of the 18 animals injected with alloxan, six did not have the running phase, whereas only one of the 17 animals given the saline solution failed to demonstrate this symptom. These frequencies could have occurred by chance between five and six times out of 100. (Fisher's exact test was used in the treatment of all frequency data.)

TABLE 1
THE EFFECT OF ALLOXAN ON THE FREQUENCY OF THE RUNNING AND CLONIC PHASES FOR THE FIRST POST-TREATMENT SEIZURE TEST

| | Running phase | | Clonic phase | |
|---|---|---|---|---|
| Group | Occurred | Did not occur | Occurred | Did not occur |
| Alloxan | 12 | 6 | 10 | 8 |
| Saline | 16 | 1 | 16 | 1 |

Those animals failing to have the running phase did not enter the clonic phase. Furthermore, two animals in the Alloxan Group, although exhibiting the running phase, did not become clonic. The frequencies of the clonic phase for the Alloxan and Saline Groups are presented in Table 1. Eight animals in the Alloxan Group did not have the clonic phase as opposed to only one animal in the Saline Group. These results are statistically significant at between the 2 and 1 per cent level.

Table 2 contains the frequencies of the running and clonic phases during the second post-treatment seizure test. These results also indicated that the alloxan treated animals had fewer running and clonic symptoms than did the saline treated animals. Statistical analyses indicate that these results could have occurred by chance less than one time out of 100.

It is of interest that every animal in the Alloxan Group failing to demonstrate the running phase during the first post-treatment seizure test did not have the running symptom on the second post-treatment seizure test. This same relationship was also observed for the clonic phase.

TABLE 2

THE EFFECT OF ALLOXAN ON THE FREQUENCY OF THE RUNNING AND CLONIC PHASES FOR THE SECOND POST-TREATMENT SEIZURE TEST

| Group | Running phase | | Clonic phase | |
|---|---|---|---|---|
| | Occurred | Did not occur | Occurred | Did not occur |
| Alloxan | 7 | 7 | 4 | 10 |
| Saline | 17 | 0 | 17 | 0 |

TABLE 3

THE EFFECT OF ALLOXAN ON THE LATENCY OF THE RUNNING AND CLONIC PHASES FOR THE FIRST POST-TREATMENT SEIZURE TEST

The mean is based on difference scores: The latency on the first post-treatment seizure test minus the latency on the pretreatment seizure test for each animal demonstrating the phase in question on both seizure tests. Latency was measured in seconds.

| | Running phase | | Clonic Phase | |
|---|---|---|---|---|
| | Saline group | Alloxan group | Saline group | Alloxan group |
| *N* | 16.00 | 10.00 | 16.00 | 10.00 |
| Mean | 0.49 | 4.19 | 0.95 | 4.80 |
| *SD* | 1.08 | 3.19 | 2.14 | 3.54 |
| *t** | | 3.36 | | 2.91 |

*As a consequence of the heterogeneity of variance, the *t*-test used in this analysis was that suggested in Edwards (2, p. 166 ff).

The effect of the experimental treatments on the latencies of the running and clonic phases for those animals still demonstrating these symptoms can be noted from an examination of Table 3. This table contains the mean difference between the latencies on the first post-treatment seizure test and those for the pretreatment seizure test.

For both the running and clonic phases, the Alloxan Group demonstrated a much greater increase in latency than did the Saline Group. The difference between the mean increase in running latency for the two groups proved to be significant at greater than the 1 per cent level. That for the clonic phase was significant at greater than the 3 per cent level. (Probabilities based on one-tailed tests.)

The latency scores for the second post-treatment seizure test were not analyzed since the number of animals was too limited.

### 2. *Experiment II*

All animals under both conditions entered into the running phase, whereas only one animal failed to demonstrate the convulsive symptom. This occurred following the administration of saline.

As a consequence of the experimental design, it was possible to take advantage of any response consistency existing for a given animal. This was accomplished by taking the difference in latencies obtained under the two different conditions. The formula used in determining this difference in latency for a given animal and a given phase was: latency following dextrose injection minus latency following saline injection.

The mean differences in latency for the running and convulsive phases are presented in Table 4. Due to one animal's not demonstrating the clonic symptom following the administration of saline, the mean difference for the clonic phase is based on 13 differences.

TABLE 4

THE EFFECT OF HIGH BLOOD SUGAR (DEXTROSE INJECTION) ON THE LATENCY OF THE RUNNING AND CLONIC PHASES

The mean is based on difference scores: The latency following the injection of dextrose minus the latency following the injection of physiological saline for each animal demonstrating the phase in question on both seizure tests. Latency was measured in seconds.

| | Phase | |
|---|---|---|
| | Running | Clonic |
| $N$ | 14.00 | 13.00 |
| Mean | —0.46 | —4.58 |
| $SE_{\text{Mean}}$ | 0.20 | 3.12 |
| $t$ | 2.28 | 1.59 |

For both the running and clonic phases, the mean difference in latency proved to be negative, indicating that, on the average, the latencies were *shorter* following the injection of dextrose.

The mean difference in latency for the running phase proved to be different from zero at between the one and five per cent level of significance. The analysis of the convulsive phase resulted in a $p$-value between 0.20 and 0.10.

## D. DISCUSSION AND SUMMARY

Experiment I was designed to test the hypothesis that one of the mechanisms responsible for the audiogenic seizure is an "excessive" secretion of insulin.

Two groups of seizure-susceptible mice were tested: one following the injection of alloxan, and the other following the injection of physiological saline. It was predicted that the alloxan would cause a decrease in seizure susceptibility.

This decrease in seizure susceptibility was found when both frequency and latency of seizures were considered.

It is of interest that both the running and clonic phases were affected by the alloxan treatment, although it appears that the latter is more readily susceptible.

Since the destruction of the beta cells increases as a consequence of the decrease in insulin, the blood sugar level, Experiment II was conducted. Its purpose was to determine if the results obtained in Experiment I could be attributed to the high blood sugar level.

Seizure susceptibility was determined when the seizure test followed either the intravenous injection of dextrose or the injection of saline. Although the frequency of seizures was not differentially affected by these two conditions, the latencies were. Following the injection of dextrose, the mice entered both phases sooner than when compared to the saline condition. These results are opposite to that expected if the decrease in susceptibility noted in Experiment I was a consequence of the increase in blood sugar.

These results also fail to agree with those reported earlier (6, 7, 11). In the first two reports, the injection of dextrose was found to have no effect on the seizure, whereas in the last report, a decrease in susceptibility was observed. Differences in strain, dextrose concentration, and the time interval between the injection and the seizure test might be possible causes for this lack of agreement.

An explanation of the results obtained in Experiment II can be derived from the "insulin" hypothesis. If one of the mechanisms underlying the audiogenic seizure is an excessive secretion of insulin, then any condition increasing the flow of insulin should result in an increase in seizure susceptibility. One such condition is an increase in the blood sugar level. Consequently, it might have been expected that following the injection of dextrose, the mice would be more susceptible to seizures rather than less.[3]

The results of these experiments thus indicate that a decrease in the insulin content of the pancreas, independent of the concomitant increase in the blood sugar level, decreases the sensitivity of mice to audiogenic seizures. Further research is necessary before the causal chain can be stated through which this effect is obtained.

---

[3] This explanation gains some support from the research of Mackay and Barbash (9). They report that several epileptic patients entered into a seizure following the introduction of glucose. However, unfortunately, the total number of patients demonstrating this was not reported.

## REFERENCES

1. DONALDSON, H. H. The Rat: Data and Reference Tables. Philadelphia: Wistar Institute, 1924.
2. EDWARDS, A. L. Experimental Design in Psychological Research. New York: Rinehart, 1950.
3. FINGER, F. W. Convulsive behavior in the rat. *Psychol. Bull.*, 1947, **44**, 201-248.
4. FRINGS, H., & FRINGS, M. Development of strains of albino mice with predictable susceptibilities to audiogenic seizures. *Science*, 1953, **117**, 283-284.
5. ———. The production of stocks of albino mice with predictable susceptibilities to audiogenic seizures. *Behavior*, 1953, **5**, 305-319.
6. GINSBURG, B., & ROBERTS, E. Glutamic acid and central nervous system activity. *Anat. Rec.*, 1951, **111**, 492-493.
7. GINSBURG, B., ROSS, S., ZAMIS, M. J., & PERKINS, A. Some effects of 1 (+) glutamic acid on sound-induced seizures in mice. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 134-141.
8. LUKENS, F. D. W. Alloxan diabetes. *Physiol. Rev.*, 1948, **28**, 304-330.
9. MACKAY, G. W. J., & BARBASH, H. The glucose tolerance curve in epilepsy. *J. Ment. Sci.*, 1931, **77**, 83-85.
10. MAIER, N. R. F. Studies of Abnormal Behavior in the Rat. I. The neurotic pattern and an analysis of the situation which produces it. New York: Harper, 1939.
11. VICARI, E. M., TRACY, A., & JONGBLOED, A. Effect of epinephrine, glucose, and certain steroids on fatal convulsive seizures in mice. *Proc. Soc. Exp. Biol. & Med.*, 1952, **80**, 47-50.

*420 Wood Street*
*Ithaca, New York*

*The Journal of Genetic Psychology*, 1957, **90**, 119-128.

# BOOKS

The *Journal of Genetic Psychology*, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy competent reviews at not less than $2 per printed page and not more than $3 per printed page, but not more than $15.00 for a single review.

*Conditions.* Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (*a*) Makes an important theoretical contribution; (*b*) consists largely of original experimental research; (*c*) has a creative or revolutionary influence in some special field or the entire field of psychology; (*d*) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

*Procedure.* If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. Just send in your review. It does not matter if the book in question has been reviewed before.

## (1955)

STONE, C. P., *Ed.* Annual Review of Psychology. Stanford: Annual Reviews, 1955. Pp. 517.

HAVIGHURST, R. J., & NEUGARTEN, B. L. American Indian and White Children. Chicago: Univ. Chicago Press, 1955. Pp. 335.

ALLPORT, F. H. Theories of Perception and the Concept of Structure. New York: Wiley, 1955. Pp. 709.

LEHNER, G. F. J., & KUBE, E. The Dynamics of Personal Adjustment. New York: Prentice-Hall, 1955. Pp. 498.

HARE, A. P., BORGATTA, E. F., & BALES, R. F., *Eds.* Small Groups. New York: Knopf, 1955. Pp. 666.

THORNDIKE, R. L., & HAGEN, E. Measurement and Evaluation in Psychology and Education. New York: Wiley, 1955. Pp. 575.

OESER, O. A., & EMERY, F. E. Social Structure and Personality in a Rural Community. New York: Macmillan, 1954. Pp. 279.

OESER, O. A., & HAMMOND, S. B. Social Structure and Personality in a City. New York: Macmillan, 1954. Pp. 344.

EATON, J. W., & WEIL, R. J. Culture and Mental Disorders. Glencoe: Free Press, 1955. Pp. 254.

MARRIOTT, M., *Ed.* Village India. Chicago: Chicago Univ. Press, 1955. Pp. 269.

ALLPORT, G. W. Becoming. New Haven: Yale Univ. Press, 1955. Pp. 106.

LAZARSFELD, P. F., & ROSENBERG, M., *Eds.* The Language of Social Research. Glencoe: Free Press, 1955. Pp. 590.

SCHAFFNER, B., *Ed.* Group Processes. New York: Macy, 1955. Pp. 334.

BARKER, R. G., & WRIGHT, H. F. Midwest and Its Children. Evanston: Row, Peterson, 1954. Pp. 532.

PANEL ON PSYCHOLOGY AND PHYSIOLOGY, COMMITTEE ON UNDERSEA WARFARE. Supplement to a Survey Report on Human Factors in Undersea Warfare. Washington: National Academy of Sciences—National Research Council, 1954. Pp. 92.

PANEL ON PSYCHOLOGY AND PHYSIOLOGY. A Survey Report on Human Factors in Undersea Warfare. Washington: Committee on Undersea Warfare, National Research Council, 1949. Pp. 541.

Abramson, H. A., *Ed.* Problems of Consciousness. New York: Macy, 1955. Pp. 180.

Roback, A. A., *Ed.* Present-Day Psychology. New York: Philosophical Library, 1955. Pp. 995.

Ilg, F. L., & Ames, L. B. Child Behavior. New York: Harper, 1955. Pp. 364.

Mead, M., & Wolfenstein, M., *Eds.* Childhood in Contemporary Cultures. Chicago: Univ. Chicago Press, 1955. Pp. 473.

Greenblatt, M., York, R. H., Brown, E. L. *et al.* From Custodial to Therapeutic Patient Care in Mental Hospitals. New York: Russell Sage Foundation, 1955. Pp. 497.

Himelhoch, J., & Fava, S. F. Sexual Behavior in American Society. New York: Norton, 1955. Pp. 446.

Darley, J. G., & Hagenah, T. Vocational Interest Measurement. Minneapolis: Univ. Minnesota Press, 1955. Pp. 279.

Farnsworth, P. R., *Ed.* Annual Review of Psychology. Stanford: Annual Reviews, Inc., 1956. Pp. 448.

## (1956)

Gardner, E. F., & Thompson, G. G. Social Relations and Morale in Small Groups. New York: Appleton-Century-Crofts, 1956. Pp. 312.

Glueck, S., & Glueck, E. Physique and Delinquency. New York: Harper, 1956. Pp. 339.

Klopfer, B. Developments in the Rorschach Technique. Yonkers-on-Hudson, New York: World Book, 1956. Pp. 828.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A Study of Thinking. New York: Wiley, 1956. Pp. 330.

Snell, G. D. (*Ed.*) Biology of the Laboratory Mouse. New York: Dover Publications, 1956. Pp. 497.

Brunswick, E. Perception and the Representative Design of Psychological Experiments. Berkeley: Univ. California Press, 1956. Pp. 154.

Ruesch, J., & Kees, W. Nonverbal Communication. Berkeley: Univ. California Press, 1956. Pp. 201.

# CRITICAL REVIEWS OF RECENT BOOKS

*The Journal of Genetic Psychology*, 1957, **90**, 121-125.

*(Allport, Gordon W. Becoming. New Haven: Yale Univ. Press, 1955. Pp. 106.)*

Reviewed by Cyril J. Adcock

The size of this book is no indication of its importance. In its few pages Allport manages to highlight some of the major deficiencies in modern psychological theory. It is a timely book both in the sense that it is greatly needed and in the sense that the psychological world is ready for it. He is not a lonely voice crying in the wilderness. Rather he puts in explicit form ideas which are agitating the minds of many people.

The major theme of the book is that man is not a passive organism whose behaviour can be described in terms of *reactions* to changing stimuli but an active being making his own contribution to the cosmic processes. These two contrasting views Allport attributes to a Lockean tradition and a Leibnitzian tradition respectively. The former he finds to be still dominant in Anglo-American psychology. Associationism, environmentalism, behaviourism, *S-R* psychology, positivism, operationism, mathematical models: all make the fundamental assumption that the intellect is a passive thing "acquiring content and structure only through the impact of sensation and the criss-cross of associations, much as a pan of sweet dough acquires tracings through the impress of a cookie cutter."

The Liebnitzian tradition was anticipated by Aristotle's doctrine of orexis and the doctrine of intention of St. Thomas Aquinas. Spinoza, Kant, Herbart, Brentano, Wundt, Husserl, Dewey, and Scheler belong to this tradition. Advocates of plural instincts such as McDougall he considers to occupy a halfway station since "instincts are after all multiple forces to which the individual is subject." It is somewhat surprising that Allport should think of instinctive forces, although "self-contained within the organism" as yet essentially external. Gestalt psychology with its dynamic cognitive forces he is prepared to accept. An internal force acting cognitively is seemingly more acceptable than one acting conatively. The explanation probably lies in Allport's stress on the active intellect. The gestalt effects are evidence of the active intellect whereas an instinct may produce action which the intellect might not approve. Some doctrines of instinct (it would be better to speak of *drives*) would merit such an attitude, but one would query its aptness in the case of McDougall whom he specifically mentions. Viewed as a whole the McDougallian system involves so many similarities with Allport's general theory that one wonders why he has not made more specific acknowledgment.

The major theme of the active being is bound up with two others: the concept of self and the notion of freedom. These are key concepts in any system which aims to get beyond mechanism. It is to be hoped that their clarification and support by such a distinguished psychologist will lead to their better recognition in current research.

Allport points out that the concept of self was thrown out of psychology with the concept of soul because its question-begging nature made it so dangerous. In recent years he finds that the tide has turned and self and ego are unashamedly back. "The situation today seems to be that many psychologists who first fit psychology to an external set of co-ordinates are dissatisfied with the result. They therefore re-invent the ego because they find no coherence among the measures yielded by positivistic analysis." This face-saving use of the concept, however, he fears may do more harm than good. He prefers a more basic concept and proposes to refer to it as the *proprium*. This includes all the aspects of personality which make for inward unity. It is these which seem warm and important, the "regions of our life which we regard as peculiarly ours." In this connection he explodes the notion that ego-involved is just emotionally charged behaviour. It is certainly true that we are more inclined to be emotional in regard to ego-involved matters but we can also be very emotional when not ego-involved.

Several aspects of the proprium are described. First we have the *bodily aspect* which serves as a permanent anchor for our self awareness. Then we have the fact of *self identity* as remarked on by William James who has contributed much to the concept of self. *Ego-enhancement* is quite distinct from ego extension which is thought of rather in terms of widening range of identification. In referring to the *rational aspect* of the proprium Allport reminds us that because rationalization occurs we should not be deceived into thinking that there is no rational thinking. The *self-image* and the *ideal self-image* are briefly referred to but much more space is given to propriate, i.e., ego-involved, *striving*. This latter is always distinguished by the fact that it makes for unification of the personality and always has a future reference. Finally the proprium is the *knower*. On this point he parts company with James but makes his point in typically Jamesian language:—

> We know not only things but we know (i.e., are acquainted with) the empirical features of our own proprium. It is I who have bodily sensations, I who recognize my self-identity from day to day; I who note and reflect on my self-assertion, self extension, my own rationalization, as well as upon my interests and striving. When I thus think about my own propriate functions I am likely to perceive their essential togetherness, and feel them bound in some way to the knowing function itself.

Allport goes on to relate the proprium to conscience and values and here again he is in opposition to any simple conditioning view. For him conscience changes in form with increased development and he suggests that we have mistaken ideas about conscience because we have largely confined our studies to children and not investigated more mature individuals. In the adult:

> 1. External sanctions give way to internal—a change adequately accounted for by the processes of identification and introjection familiar in Freudian and behavioural theory.
> 2. Experiences of prohibition, fear and "must" give way to experience of preference, self-respect and "ought." This shift becomes possible in proportion as the self-image and value-systems of the individual develop.
> 3. Specific habits of obedience give way to generic self-guidance, that is to say, to broad schemata of values that confer direction upon conduct.

Most of this has been said before. McDougall, in particular, has said it with much more detail. He too has described how conscience rises above the factors which gave rise to it. He too, despite his plural instincts, accepts an emergent view of conscience. The value of Allport's work here lies not in its originality but in his simple, clear exposition and his relation of it to the general theme of the Leibnitzian concept of man as an active being rather than a passive reactor, a dynamic psychological unity rather than a bundle of independent drives, in no sense a machine no matter how complicated.

This brings us to the third theme, that of freedom. It is certainly true that we can look through a "hundred successive American" books on psychology and find no mention of "will" or "freedom." As good scientists we must regard freedom as an illusion, at least volitional freedom. On the other hand commonsense has always favoured the notion of free will. Allport comes out on the side of commonsense. He first makes a distinction between the reference frame of the scientist and the behaving individual. The former may foresee the probability of certain behaviour because from his wider viewpoint he can see the circumstances which will require it. Secondly, he points out that certain conditions make for *relatively* more or less freedom for the individual, e.g., self insight. Thirdly, relative freedom depends on the possession of multiple possibilities for behaviour. *More* determining tendencies in the individual means *more* freedom. Finally, as psychologists, we know that there is greater freedom in some forms of response than others, e.g., calm reasoning as opposed to blind rage.

There are two essential points in all this reasoning and they may be better stated more explicitly. The first is that freedom can have no meaning to the individual except in terms of his own value systems. There is abroad in some quarters the quaint notion that freedom of will can be manifested only in the possibility of the individual doing that which is contra-indicated by his value systems. This would be not freedom but madness. Its basis is possibly the regarding of the will as some kind of homunculus. Will which is not identified with the person is meaningless. Freedom for the will is nothing more than the possibility of self determination and this is a perfectly respectable scientific concept.

The second point is that self-determination may be possible in varying degrees. Behaviour is always to be understood in terms of the factors lying within the individual and his environment. The more of these factors lie within the individual relative to those in the environment the greater is the freedom of the individual. In this sense complete freedom requires divinity.

Allport might have gone on to point out that when we regard freedom as a positive thing and not merely in the negative sense of absence of coercion we find another paradox—that increase of laws to be obeyed brings increase of freedom. The citizen of New York with all his irksome restrictions is in this sense much freer than Robinson Crusoe who was monarch of all he surveyed. By agreeing to restrict my driving to the right side of the road I attain greater freedom than if I insisted on my freedom to drive where I wish.

The question of freedom is one which appears in its true significance only when we recognize the essential unity of the person as focussed in the ego or proprium. In the degree to which Allport has brought out this aspect of human behaviour he has done an important service to psychology. He has made an important contribution to the humanizing of psychology. The growing importance of the so-called *cognitive school* and the *transactional approach* are other indications of a developing trend. But it is not new schools we need. Our aim must be an all-embracing theory which reconciles all the differences of the schools, and of this too there have been some indications recently.

In stressing the three major themes in this book we must be careful not to imply that there is not a wealth of other material. Allport is still interested in the *unique personality* but he is inclined to give more importance to the inherited components in this uniqueness. The work of Tinbergen and his school have had some influence here. But we must not continue or we shall have presented the whole of Allport's book which would lose much by second hand presentation. The reader is advised to make sure of reading the original. No matter how much he may agree or disagree with it he can hardly fail to enjoy it.

*Department of Psychology*
*Victoria University College*
*P. O. Box 196*
*Wellington, New Zealand*

# BOOKS RECENTLY RECEIVED

*(There will always be two pages of book titles, listed in the order of receipt, i.e., the most recently received books will be found at the end of the list.)*

Darley, J. G., & Hagenah, T. Vocational Interest Measurement. Minneapolis: Univ. Minnesota Press, 1955. Pp. 279.

Strauss, A., *Ed.* The Social Psychology of George Herbert Mead. Chicago: Univ. Chicago Press, 1956. Pp. 296.

United Nations. Training for Social Work. New York: Columbia Univ. Press, 1956. Pp. 160.

Pike, K. L. Language. Glendale: Summer Institute of Linguistics, 1955. Pp. 85.

Kwalwasser, J. Exploring the Musical Mind. Boston: Coleman-Ross, 1955. Pp. 189.

Stogdill, R. M., & Shartle, C. L. Methods in the Study of Administrative Leadership. Columbus Ohio State Univ., 1955. Pp. 77.

Strasser, S. Das Gemut. Antwerp: Spectrum/Herder, 1956. Pp. 291.

Rees, J. T., & Usill, H. V., *Eds.* They Stand Apart. New York: Macmillan, 1955. Pp. 220.

Farnsworth, P. R., *Ed.* Annual Review of Psychology. Stanford: Annual Reviews, Inc., 1958. Pp. 448.

Ruesch, J., & Kees, W. Nonverbal Communication. Berkeley: Univ. California Press, 1956. Pp. 201.

Brunswick, E. Perception and the Representative Design of Psychological Experiments. Berkeley: Univ. California Press, 1956. Pp. 154.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A Study of Thinking. New York: Wiley, 1956. Pp. 330.

Myrdal, A., & Klein, V. Women's Two Rôles. New York: Humanities Press, 1956. Pp. 208.

Benney, M., Gray, A. P., & Pear, R. H. How People Vote. New York: Humanities Press, 1956. Pp. 227.

Klein, J. The Study of Groups. New York: Humanities Press, 1956. Pp. 200.

Crow, L. D., & Crow, A. Understanding Our Behavior. New York: Knopf, 1956. Pp. 347.

Klein, D. B. Mental Hygiene. New York: Holt, 1956. Pp. 654.

Ewing, U. C. Thresholds of Existence. New York: Philosophical Library, 1956. Pp. 286.

Gray, W. S., & Rogers, B. Maturity in Reading. Chicago: Univ. Chicago Press, 1956. Pp. 273.

Mayer, F., & Brower, F. E. Education for Maturity. Washington, D. C.: Public Affairs Press, 1956. Pp. 155.

Stephens, J. M. Educational Psychology. New York: Holt, 1956. Pp. 717.

Seeley, J. R., Sim, R. A., & Looseley, E. W. Crestwood Heights. New York: Basic Books, 1956. Pp. 505.

Schwartz, M. S., & Shockley, E. L. The Nurse and the Mental Patient. New York: Russell Sage Foundation, 1956. Pp. 289.

Glueck, S., & Glueck, E. Physique and Delinquency. New York: Harper, 1956. Pp. 339.

Roe, A. The Psychology of Occupations. New York: Wiley, 1956. Pp. 340.

Sartre, J. P. Being and Nothingness. New York: Philosophical Library, 1956. Pp. 638.

Snell, G. D. Biology of the Laboratory Mouse. New York: Dover Publications, 1956. Pp. 497.

Klopfer, B. Developments in the Rorschach Technique. Yonkers-on-Hudson, New York: World Book, 1956. Pp. 828.

Crow, L. D., & Crow, A. Adolescent Development and Adjustment. New York: McGraw-Hill, 1956. Pp. 555.

Rytand, D. A., & Creger, W., *Eds.* Annual Review of Medicine. Stanford, California: Annual Reviews, 1956. Pp. 611.

Borgatta, E. F., & Meyer, H. J. Sociological Theory. New York: Knopf, 1956. Pp. 546.

Brameld, T. Toward a Reconstructed Philosophy of Education. New York: Dryden, 1956. Pp. 417.

Stevenson, G. S. Mental Health Planning for Social Action. New York: McGraw-Hill, The Blakiston Division, 1956. Pp. 358.

Wenger, M. A., Jones, F. N., & Jones, M. H. Physiological Psychology. New New York: Holt, 1956. Pp. 472.

Du Mas, F. M. Manifest Structure Analysis. Missoula, Montana: Montana State Univ. Press, 1955. Pp. 193.

Mead, M. New Lives for Old. New York: Morrow, 1956. Pp. 548.

Gardner, E. F., & Thompson, G. G. Social Relations and Morale in Small Groups. New York: Appleton-Century-Crofts, 1956. Pp. 312.

Johnson, R. C. Psychical Research. New York: Philosophical Library, 1956. Pp. 176.

Henry, W. E. The Analysis of Fantasy. New York: Wiley, 1956. Pp. 305.

Morris, C. Varieties of Human Value. Chicago: Univ. Chicago Press, 1956. Pp. 208.

Olds, J. The Growth and Structure of Motives. Glencoe, Illinois: Free Press, 1956. Pp. 277.

Marzolf, S. S. Psychological Diagnosis and Counseling in the Schools. New York: Holt, 1956. Pp. 401.

Lindesmith, A. R., & Strauss, A. L. Social Psychology. New York: Dryden Press, 1956. Pp. 703.

Abramson, H. A., *Ed.* Neuropharmacology. New York: Josiah Macy, Jr. Foundation, 1956. Pp. 328.

Bossard, J H. S., & Boll, E. S. The Large Family System. Philadelphia: Univ. Pennsylvania Press, 1956. Pp. 325.

Broudy, H. S., & Freel, E. L. Psychology for General Education. New York: Longmans, Green, 1956. Pp. 456.

Rose, A. M. Sociology: The Study of Human Relations. New York: Knopf, 1956. Pp. 589.

Festinger, L., Riecken, H. W., & Schachter, S. When Prophecy Fails. Minneapolis: Univ. Minnesota Press, 1956. Pp. 256.

Schaffer, L. F., & Shoben, E. J. The Psychology of Adjustment. Boston: Houghton Mifflin, 1956. Pp. 672.

Bureau of Publications, Teachers College. Communication and the Communication Arts. New York: Columbia University, 1956. Pp. 149.

Greenwood, D. Essays in Human Relations. Washington, D. C.: Public Affairs Press, 1956. Pp. 76.

Simmons, H. E. The Psychosomatic Aspects of Cancer. Washington, D. C.: Peabody Press, 1956. Pp. 53.

Barton, R. C. Our Human Rights. Washington, D. C.: Public Affairs Press, 1955. Pp. 102.

Jordan, A. M. Educational Psychology. New York: Holt, 1956. Pp. 600.

American Friends Service Committee. Indians of California, Past and Present. 1830 Sutter St., San Francisco, 1956. Pp. 36.

McEathron, M. Your Child Can Learn to Read. New York: Grosset & Dunlap, 1952. Pp. 92.

Flesch, R. Teaching Johnny to Read. New York: Grosset & Dunlap, 1956. Pp. 92.

Burnham, J. C. Lester Frank Ward in American Thought. Washington, D. C.: Public Affairs Press, 1956. Pp. 31.

Peiper, A. Die Eigenart der Kindlichen Hirntatigkeit. Leipzig: Veb Georg Thieme, 1956. Pp. 588.

Kainz, F. Psychologie der Sprache. Stuttgart: Ferdinand Enke Verlag, 1956. Pp. 537.

Centre National de la Recherche Scientifique. L'Analyse Factorielle et Ses Applications. Paris: France, 1955. Pp. 431.

Bjerstedt, A. Interpretations of Sociometric Choice Status. Copenhagen: Munksgaard, 1956. Pp. 408.

Stewart, D. A. Preface to Empathy. New York: Philosophical Library, 1956. Pp. 157.

Hankin, F. Making Democracy Work. Washington, D. C.: Public Affairs Press, 1956. Pp. 74.

$14.00 per year
$10.00 per volume
Single Numbers $5.00

QUARTERLY
Two volumes per year

June, 1957
Volume 90, Second Half

Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

# THE JOURNAL OF GENETIC PSYCHOLOGY

## Child Behavior, Animal Behavior, and Comparative Psychology

---

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

## JUNE, 1957

**(Manuscripts are printed in the order of final acceptance)**

The relationship between intelligence and age among superior adults . . . 131
By Edwin E. Ghiselli

Word variety as a measure of bilingualism in preschool children . . . 143
By Madorah E. Smith

Relationships between early social experience and adult aggressive behavior in inbred mice . . . . . . . . . . . . . 151
By John A. King

Some hypotheses on perceptual and learning processes with their applications to the process of reading: A preliminary note . . . . . . . 167
By Helene Powner Mann

The effect of frontal cortical lesions on activity wheel and open-field behavior 203
By John A. Stern

(OVER)

---

Copyright, 1957, by The Journal Press
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

Hoarding behavior in the hooded rat as a function of varied alley illumination . . . . . . . . . . . . . . . . . 213
By Melvin H. Marx, Shinkuro Iwahara, and Aaron J. Brownstein

Spread of effect: Animal studies . . . . . . . . . . . . 219
By Melvin H. Marx

The distinction between living and not living among 7-10-year-old children, with some remarks concerning the so-called animism controversy . . 227
By Göte Klingberg

Academic achievement in a group of mentally disturbed adolescents in a residential treatment setting . . . . . . . . . . 239
By Kurt Salzinger

Progress in the use of English after 22 years by children of Chinese ancestry in Honolulu . . . . . . . . . . . . 255
By Madorah E. Smith

Latent extinction in a multiple-*T* Maze within heterogeneous and homogeneous environments . . . . . . . . . . . . . 259
By W. Lynn Brown and Edward S. Halas

SHORT ARTICLES AND NOTES . . . . . . . . . . 267

Drive in illness: An experimental study of barrier behavior of (alloxan) diabetic and nondiabetic animals (William Seeman and Robert C. Green, Jr.) . . . . . . . . . . . . . 267

Graduate schools ought to train teachers of psychology (Virginia M. Staudt) . . . . . . . . . . . . . . 271

BOOKS . . . . . . . . . . . . . . . 275

BOOKS RECENTLY PUBLISHED . . . . . . . . . 275

Henry, W. E. *The Analysis of Fantasy: The Thematic Apperception Technique in the Study of Personality*. (Reviewed by Robert M. Allen) . . . . . . . . . . . . . 280

*The Journal of Genetic Psychology*, 1957, **90**, 131-142.

# THE RELATIONSHIP BETWEEN INTELLIGENCE AND AGE AMONG SUPERIOR ADULTS*

*Department of Psychology, University of California, Berkeley*

EDWIN E. GHISELLI

## A. INTRODUCTION

The relationship between intelligence and age at the adult level has been the subject of considerable discussion and numerous investigations. The nature of this relationship not only is of theoretical interest, but also it obviously has many practical implications. Knowledge of the effects of age upon higher mental processes casts further light upon the nature of such processes and at the same time is useful in planning for effective utilization of manpower.

Summaries of the many studies of intelligence and age point to the striking agreement with respect to the general findings (e.g., 5, 12). As age increases from about 20 years there is a gradual and continuous decline in intelligence test score. A wide variety of measures of intellectual ability has been employed, and while the rate of decline is less with some tests than with others, the downward trend uniformly is found. There is some disagreement with respect to the shape of this curve of decrement. The most common notion is that with increasing years there is an acceleration in the rate of decline. The amount of drop in average score during the early years past maturity is thought to be considerably less than the amount of drop manifest during later years. However, Vincent, on the basis of an analysis of a number of studies, concluded that for all practical purposes the relationship can be considered to be linear (12). It therefore appears that from maturity on each additional year takes equal toll.

It is true that all abilities do not show the same kind of decline that is manifest with intelligence tests. Indeed, on some types of tests the average scores earned by older persons actually are higher than those of younger persons. Thus Brown and Ghiselli report an increase in arithmetic test score of about half a standard deviation from 20 to 50 years of age (2). The change in performance with age, then, is in part a function of the nature of the ability measured and the type of test employed.

Since there is some variation in the content and nature of different intel-

---

*Received in the Editorial Office on January 3, 1956.

Bureau Ednl. Psy. Research
DAVID HA ING COLLEGE
Dated
Accs. No

ligence tests, it is not surprising to find that variations in the rate of decrement in intelligence test scores with age have been found with different types of intelligence tests and with different parts of the same test. In their summary of pertinent studies, Jones and Kaplan conclude that older adults are most handicapped in tests wherein speed is an important factor (5). Lorge, for example, using a variety of intelligence tests showed that the decline in scores is greater with speed tests than with tests wherein unlimited time is permitted and the individual can set his own pace in answering the questions (7). Jones and Conrad found that there is likely to be less decline in scores with tests that place a premium on cumulative experience (4). These effects of speed and familiarity of material have been demonstrated by Wechsler (13), and Brown and Ghiselli (2).

There is some suggestion that the reduction in intelligence test scores is less extensive at the higher levels of ability than at the lower levels (8, 11). The effect of ability level is most strikingly demonstrated by the results obtained by Owens (9). He administered a speeded intelligence test to a group of college freshmen and then retested them some 30 years later. On the average total test scores increased about half a standard deviation from the first to the second administration. The greatest increases were with material that might well be influenced by cumulative experience, and there were little or no changes in performance on material where speed is of particular importance. In addition, with material measuring higher mental process (e.g., analogous reasoning) there tended to be an improvement in performance whereas with simpler material (e.g., following directions) there was little or no change in performance.

From the foregoing review it seems clear that in general there is a decline in intelligence test scores of adults with increasing age. The decrease is least when speed is not a factor, when the material is subject to the effects of cumulative experience, and when the problems involve the higher mental processes. In addition, with persons at the higher levels of ability the effects of age are less severe. Finally, there is the suggestion that under optimal conditions there may be no decrease in intelligence test score at all.

It is the purpose of the present study to examine the relationship between intelligence test scores and age of adults under optimal conditions. The hypothesis is advanced that the performance of superior adults on an intellectual task of a very high level without the pressure of time will show little or no deterioration with age.

No matter what an intelligence test may be titled, the fact that it is designed to measure higher abilities is perfectly apparent to the individual who takes it. Whether the test is called the *X* Measure of Mental Alertness, the *Y* Aptitude Index, or merely the *Z* test, the very first question reveals the intent of the test. To a younger person, particularly a college student, an intelligence test simply may be an intellectual challenge. But an older person may view such a test not only with suspicion but also with fear. A device that measures intelligence but is made up of material quite different from intelligence test items may be reacted to more favorably. If the lower scores earned by older persons are the result of such attitudinal or cultural factors as Kuhlen (6) suggests, then such a "disguised" measure of intelligence should be even less likely to show a decrement in scores than one the objective of which is obvious.

## B. Measuring Instruments

Two devices were used to obtain measurements of intellectual ability. One entitled "Analysis of Relationships" was used to give a direct measure of intelligence, and the other entitled "Self-Description Inventory," was used to give a "disguised" or indirect measure of the trait.

### 1. *Analysis of Relationships Test*

The Analysis of Relationships Test was developed to measure the same general types of general intellectual abilities as ordinarily are measured by intelligence tests. In order to get at the so-called higher mental processes, the attempt was made to give primary emphasis to items of reasoning and problem solving. The test contains five types of items. Eleven involve quantitative reasoning, nine vocabulary (largely through similarities and opposites), 10 number series extension, 12 analogies, and eight proverbs. Subjects were allowed to take all the time they needed to complete the test, and the average time taken was of the order of 40 minutes.

Reliability of the test was determined from the coefficient of correlation between odd and even items, adjusted by the Spearman-Brown formula. Data on the reliability of the test for several different groups are given in Table 1. It will be observed that the reliability of the instrument is satisfactory for a wide variety of individuals. The lowest reliability coefficient is .77 and the highest is .90.

In order to ascertain the extent of sex differences, 44 employed men were matched with 44 women for Otis Employment Test score. The average score of the men was found to be .3 points higher than that of women. In addition 49 senior high school boys were matched with a like number of girls for *IQ*. The average score of the boys was .9 points higher than that of girls.

It therefore appears that the scores of males are slightly superior to those of females, though probably insignificantly so.

TABLE 1
RELIABILITY OF THE ANALYSIS OF RELATIONSHIPS TEST FOR VARIOUS DIFFERENT GROUPS OF INDIVIDUALS

| Reliability coefficient | No. of cases | Group |
|---|---|---|
| .90 | 123 | Graduating high school seniors |
| .83 | 290 | University freshmen and sophomores |
| .82 | 402 | University juniors and seniors |
| .77 | 81 | Executive, administrative, and staff personnel |
| .77 | 43 | First line supervisors |
| .81 | 50 | Office workers |
| .88 | 36 | Technical and skilled workers |

There certainly is little agreement as to how demonstration can be made of the validity of an intelligence test. In general, to be accepted as a measure of intelligence, scores on a test need to be highly related to scores on other intelligence tests, and positively correlated to at least a moderate degree with scores on tests of other types of intellectual ability, school grades, and success in higher level occupations. Table 2 gives the coefficients of correlation between scores on the Analysis of Relationships Test and other measures of general intelligence and tests of other intellectual abilities. Scores on the Analysis of Relationships Test bear high correlations with scores on other intelligence tests, and moderate correlations with scores on tests of other intellectual abilities. Table 3 gives the coefficients of correlation with school grades. These validity coefficients are of the general order of .50 which is about the value usually reported for measures of general intelligence. The coefficients of correlation between scores and indices of success on higher jobs given in Table 4 are reasonably substantial, ranging from .22 to .76. These validity coefficients are fully as high as, or even higher than, those generally reported. Therefore the criteria of validity listed at the beginning of this paragraph appear to be well satisfied by the Analysis of Relationships Test.

## 2. *Self-Description Inventory*

The details of the Self-Description Inventory have been described elsewhere (3). The inventory consists of 64 pairs of descriptive adjectives equated for social desirability. The first 32 pairs consist of favorable adjectives, and the last 32 are unfavorable ones. For the first 32 items the subjects are instructed as follows: "In each of the pairs of words below check the one you think *most* describes you," and for the last 32 items the word *least* is substituted for *most*.

TABLE 2
COEFFICIENTS OF CORRELATION BETWEEN SCORES ON ANALYSIS OF RELATIONSHIPS TEST AND SCORES ON OTHER TESTS

| Coefficient of correlation | Variable correlated | No. of cases | Group |
|---|---|---|---|
| .70 | *IQ* | 102 | Graduating high school seniors |
| .78 | Otis Employment test | 127 | Employed persons |
| .37 | Vocabulary | 69 | University sophomores |
| .35 | Vocabulary | 36 | Employed persons |
| .39 | Arithmetic speed and accuracy | 63 | Employed persons |
| .58 | Mechanical principles | 46 | Employed men |

TABLE 3
COEFFICIENTS OF CORRELATION BETWEEN SCORES ON THE ANALYSIS OF RELATIONSHIPS TEST AND SCHOOL GRADES

| Coefficient of correlation | Variable correlated | No. of cases | Group |
|---|---|---|---|
| .65 | 4-year high school grade point average | 118 | Graduating high school seniors |
| .25 | First semester grades | 37 | University freshmen |
| .59 | First year grades | 100 | University sophomores |
| .65 | First 2-year grades | 213 | University juniors |
| .32 | First 3-year grades | 78 | University seniors |

TABLE 4
COEFFICIENTS OF CORRELATION BETWEEN SCORES ON THE ANALYSIS OF RELATIONSHIPS TEST AND INDICES OF SUCCESS IN HIGHER LEVEL JOBS

| Coefficient of correlation | Variable correlated | No. of cases | Group |
|---|---|---|---|
| .64 | No. of clerks supervised | 31 | Office supervisors in insurance companies |
| .30 | Ratings of job proficiency | 24 | Supervisors in different industrial organizations |
| .31 | No. of workers supervised | 24 | Supervisors in different industrial organizations |
| .22 | Ratings of job proficiency | 39 | Personnel officers in insurance companies |
| .27 | Ratings of job proficiency | 22 | Technical and skilled workers |
| .63 | Ratings of job proficiency | 39 | Office workers |
| .76 | Ratings of job proficiency | 22 | Supervisors in a food processing company |

In order to develop a "disguised" or indirect measure of intelligence, the Analysis of Relationships Test and the Self-Description Inventory were administered to 296 employed persons. About 70 per cent of these people were men and 30 per cent women. They varied in age from 20 to 60, and occupationally from unskilled workers to top management people. An item analysis was performed on the Self-Description Inventory using the intelligence test as the criterion. Of the 64 items, 33 were found to yield responses significantly related to the intelligence test scores. These items were selected for the final scale and were weighted in terms of their correlation with the criterion test.

In order to cross-validate this "intelligence" scale, scores on the scale were related to scores on the Analysis of Relationships Test using different groups, to scores on the Otis Employment Test, and to grades of undergraduate university students. The coefficients of correlation are given in Table 5. By and large these coefficients of correlation are moderately high, being of the order of about .55 for the intelligence test. These values are about as high as can be expected from so short a scale, and are comparable with the validity coefficient of other so-called non-intellectual predictors of intelligence.

TABLE 5

COEFFICIENTS OF CORRELATION BETWEEN SCORES ON THE INTELLIGENCE SCALE OF THE SELF-DESCRIPTION INVENTORY AND SCORES ON THE VARIOUS OTHER VARIABLES

| Coefficient of correlation | Variable correlated | No. of cases | Group |
|---|---|---|---|
| .58 | Analysis of Relationships | 43 | College Students |
| .60 | Analysis of Relationships | 20 | Employed Persons |
| .52 | Otis Employment Test | 148 | Employed Persons |
| .38 | Grades | 114 | Undergraduate Students |

Thus the Self-Description Inventory appears at least reasonably satisfactory for use as a non-intellectual measure of intelligence. It yields an index of intelligence under circumstances wherein the subjects are presented with a task not obviously related to intellectual problems. Indeed, in administering the inventory subjects were informed that the inventory measured aspects of personality rather than ability, and that there were no right or wrong answers.

## C. THE SAMPLES OF INDIVIDUALS

On the Analysis of Relationships Test scores were collected for 628 adults, and on the Self-Description Inventory for 795 adults. For both instruments about 70 per cent of the subjects were men and 30 per cent women. Most of the younger subjects were college students who took the tests either as a demonstration in a course in psychology, or in connection with some experimental investigation. The employed persons were drawn from a variety of sources. Some of them were enrolled in evening schools, some were tested relative to employment or upgrading, a number were tested in connection with personnel surveys in their organizations, and some were volunteers. Only those cases were used who reported having completed one or more years of college work. The lowest occupation represented was skilled worker, and the few in this category were largely highly proficient technicians. The largest bulk were clerical, sales, and managerial personnel.

## D. RESULTS

The distributions of scores by age on the Analysis of Relationships Test and the intelligence scale of the Self-Description Inventory are given in Table 6 and 7, respectively. The means and standard deviations of these distributions are presented in Tables 8 and 9.

The raw scores for the Analysis of Relationships Test show no consistent trend either upward or downward with age. These results can be taken to indicate that there is no systematic change in average test performance throughout the age range studied. The picture is not quite so clear with the standard deviations of the scores of the various age groups. However, the indication is that there is a decrease in variability with increasing age. Homogeneity is greatest with the very oldest groups.

The results obtained with the Self-Description Inventory are quite consistent with those obtained using the Analysis of Relationships Test. The mean values fluctuate irregularly showing no consistent trend either upward or downward. A best fitting line through the points would be a horizontal straight line. With increasing age there again is a tendency for the standard deviations to become smaller in magnitude.

From these two sets of results, then, it can be concluded that with these two measures of intellectual ability there is no change in average score during the span of years from 20 to 60. The ability of older persons is equal to that of younger ones. However, with increasing age there is some decrease in variability. Proportionately fewer older persons make either higher or lower scores than do younger persons. Older persons, therefore, are more like each other in ability than are younger persons.

## E. DISCUSSION

The hypothesis the present investigation was designed to examine stated that given optimal testing conditions with superior adults there is little or no decrement in average intelligence test score with increasing age. The results obtained with both measures of intelligence here employed support this hypothesis. It therefore appears that the general reduction in intelligence with age so uniformly reported in the literature does not hold with all individuals and with all devices for measuring intelligence. When the psychological functions measured are very high mental processes, there is no pressure of time, and the persons involved are superior adults, there is no change in ability from maturity to 60 years of age.

TABLE 6
DISTRIBUTION OF SCORES ON THE ANALYSIS OF RELATIONSHIPS TEST BY AGE

| | Age in years | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Raw score | 20 to 24 | 25 to 29 | 30 to 34 | 35 to 39 | 40 to 44 | 45 to 49 | 50 to 54 | 55 to 64 |
| 50 | | | | 1 | | | | |
| 49 | 1 | | | | | | | |
| 48 | 2 | | 2 | | | | | |
| 47 | 4 | 1 | 1 | 1 | | | | |
| 46 | 9 | 2 | 2 | 1 | 1 | | | |
| 45 | 8 | 6 | 4 | 2 | | 1 | | |
| 44 | 13 | 4 | 4 | 2 | 3 | | | |
| 43 | 16 | 9 | 5 | 2 | 3 | 1 | 1 | |
| 42 | 17 | 9 | 2 | 5 | 3 | 4 | | 1 |
| 41 | 17 | 9 | 5 | 2 | | 3 | 2 | |
| 40 | 26 | 5 | 2 | 7 | 2 | | 2 | 1 |
| 39 | 23 | 12 | 3 | 5 | 3 | 1 | 2 | |
| 38 | 26 | 12 | 4 | 5 | 2 | | 1 | 1 |
| 37 | 23 | 10 | 5 | 3 | 3 | 2 | | |
| 36 | 21 | 5 | 2 | 1 | 2 | 2 | | 2 |
| 35 | 20 | 8 | 4 | 5 | 2 | 1 | | 1 |
| 34 | 8 | 3 | 5 | 3 | | 1 | | |
| 33 | 15 | 7 | 3 | 1 | 2 | 1 | | |
| 32 | 3 | 4 | 5 | 3 | 3 | | | |
| 31 | 8 | 3 | 5 | 1 | | 1 | | |
| 30 | 12 | 3 | 5 | 1 | | 1 | | |
| 29 | 13 | 3 | 3 | | 2 | | | |
| 28 | 4 | 2 | | 1 | 1 | 1 | | |
| 27 | 6 | 1 | | | | | | |
| 26 | 3 | | | 1 | | | | |
| 25 | 2 | 1 | | | | | | |
| 24 | 1 | 1 | 2 | | | | | |
| 23 | | | | | | | | |
| 22 | 1 | 1 | | | 1 | | | |
| 21 | | 1 | | | 2 | | | |
| 20 | 1 | | | | 1 | | | |
| 19 | 1 | 1 | | | | | | |
| 18 | 1 | | | | | | | |
| 17 | 1 | | | | | 2 | | |
| 16 | | | | | 1 | | | |
| 15 | | | | 1 | | | | |

Since with the tests used here no change in average performance with age was manifest, one might be tempted to conclude that there are no changes in the underlying abilities and traits measured. Such a conclusion, however,

TABLE 7
DISTRIBUTION OF SCORES ON THE INTELLIGENCE SCALE OF THE SELF DESCRIPTION INVENTORY BY AGE

| | Age in years | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Raw score | 20 to 24 | 25 to 29 | 30 to 34 | 35 to 39 | 40 to 44 | 45 to 49 | 50 to 54 | 55 to 64 |
| 63 | | | 1 | | | | | |
| 62 | | | | | | | | |
| 61 | | | | | | | | |
| 60 | | | | | | | | |
| 59 | 2 | 1 | | | | | | |
| 58 | | 1 | 1 | | | | | |
| 57 | 1 | 3 | | | | | | |
| 56 | 3 | 2 | | | | | | |
| 55 | 4 | 3 | 1 | 1 | | | | |
| 54 | 5 | 2 | | | | 1 | | |
| 53 | 7 | 3 | 3 | 2 | 2 | | 2 | |
| 52 | 7 | 4 | 3 | 1 | 2 | 1 | | |
| 51 | 10 | | 1 | 5 | | 1 | | |
| 50 | 15 | 4 | 7 | | 2 | 1 | 1 | |
| 49 | 12 | 6 | 1 | 4 | 1 | 1 | | |
| 48 | 17 | 4 | 5 | 3 | 3 | | | |
| 47 | 10 | 9 | 3 | | 2 | 4 | | |
| 46 | 11 | 7 | 3 | 2 | | | | |
| 45 | 18 | 6 | 2 | 5 | 1 | 2 | 2 | |
| 44 | 25 | 10 | 7 | 1 | 4 | | 1 | |
| 43 | 23 | 3 | 3 | 2 | 3 | | 2 | 1 |
| 42 | 17 | 18 | 5 | 6 | 5 | | | |
| 41 | 18 | 6 | 4 | 5 | 1 | 1 | 1 | |
| 40 | 18 | 7 | 2 | 5 | 5 | 2 | 1 | |
| 39 | 18 | 4 | 2 | 3 | 2 | 1 | | |
| 38 | 21 | 6 | 3 | 6 | 2 | | | |
| 37 | 16 | 10 | 4 | 4 | 1 | 2 | | |
| 36 | 17 | 4 | 2 | 5 | 4 | 1 | | |
| 35 | 10 | 3 | 5 | 1 | 3 | | | 1 |
| 34 | 16 | 5 | 1 | 4 | 3 | | | |
| 33 | 15 | 5 | 2 | 1 | | | | 2 |
| 32 | 10 | 6 | 1 | 2 | 1 | | | |
| 31 | 12 | 3 | 3 | | | | | |
| 30 | 9 | 4 | 1 | | 1 | | | |
| 29 | | 1 | 3 | | 2 | 1 | | |
| 28 | 9 | 1 | 2 | 1 | | | | 1 |
| 27 | 8 | 2 | | 1 | | | | |
| 26 | 5 | | 1 | | 1 | 1 | | |
| 25 | 1 | 1 | 1 | | | | | |
| 24 | 3 | | | 1 | | | | |
| 23 | 2 | | | 1 | | | | |
| 22 | | 1 | | | | | | |
| 21 | 2 | | 1 | | | | | |
| 20 | | | | | | | | |
| 19 | | | | | | | | |
| 18 | | | | | | | | |
| 17 | | | 1 | | | | | |
| 16 | | | | | | 1 | | |

would be most tenuous. In this connection it can be noted that Balinsky, who performed multiple factor analyses of the subtests of the Wechsler-Bellevue scale for various age groups, found some, though not great, variation in the factorial content of the measures (1). Since the two tests employed in the present study yielded only total scores, no statements can be made concerning the extent to which the functions they measure change with age. Some incidental support for the notion that there are changes, however, is obtained from an analysis of the items of the intelligence scale of the Self-Description Inventory. For Item 8 of the inventory, "determined-courageous," the first choice was scored as correct. That is, in the original development of the intelligence scale, more people scoring high on the criterion intelligence test were found to describe themselves as being "determined" rather than as "courageous," with the opposite holding for persons scoring low on the criterion intelligence test. Responses on this item also are related to age. With increasing age more and more individuals tend to describe themselves as "courageous" rather than "determined." Another example is provided by Item 13, "kind-jolly." The correct answer

TABLE 8
MEANS AND STANDARD DEVIATIONS OF SCORES ON THE ANALYSIS OF RELATIONSHIPS TEST BY AGE GROUP

| Age in years | Mean score | Standard deviation of scores | No. of cases |
|---|---|---|---|
| 20-24 | 37.18 | 5.69 | 306 |
| 25-29 | 37.38 | 5.51 | 123 |
| 30-34 | 37.30 | 5.83 | 71 |
| 35-39 | 37.91 | 5.71 | 54 |
| 40-44 | 35.61 | 7.68 | 38 |
| 45-49 | 35.86 | 7.44 | 22 |
| 50-54 | 40.13 | 1.44 | 8 |
| 54-65 | 37.83 | 2.48 | 6 |

TABLE 9
MEANS AND STANDARD DEVIATIONS OF SCORES ON THE SELF-DESCRIPTION INVENTORY BY AGE GROUP

| Age in years | Mean score | Standard deviation of scores | No. of cases |
|---|---|---|---|
| 20-24 | 40.60 | 7.64 | 397 |
| 25-29 | 41.89 | 7.53 | 155 |
| 30-34 | 41.35 | 8.59 | 85 |
| 35-39 | 41.04 | 6.83 | 72 |
| 40-44 | 40.90 | 6.47 | 51 |
| 45-49 | 32.95 | 7.37 | 21 |
| 50-54 | 45.70 | 4.45 | 10 |
| 54-65 | 34.40 | 4.88 | 5 |

on the intelligence scale is "kind." With increasing age there is an increase in the tendency to choose this alternative. Altogether there are some eight items of the intelligence scale of the Self-Description Inventory which show important relationships between type of response and age. It is evident that these changes counterbalance so that total scores remain unchanged from one age group to another. Nevertheless it is apparent that this means the functions being measured are not exactly the same at all ages.

However, from results obtained by Owens it does appear that there is quite a high degree of consistency in intelligence test scores over long periods of time during adult life (10). The coefficient of correlation between the scores his subjects earned when they were college freshmen and the scores they earned on the same test 30 years later is .77. Considering the narrow range of talent characteristic of his subjects, this coefficient is remarkably high.

## F. Summary

Previous investigations of the relationship between intelligence test scores and age at the adult level have found that from maturity on there is a steady decline in ability. An analysis of the results obtained with different types of tests, testing conditions, and subjects, suggested the hypothesis that with a test measuring the higher intellectual functions and putting no premium on speed of response, adults who have achieved higher levels of education would show little or no decline in scores with increasing age. To test this hypothesis two tests designed to measure the higher aspects of intelligence were developed. These tests were administered to groups of 628 and 795 individuals ranging in age from 20 to 65 years who had completed at least one year of college. On neither test was there found a downward trend in scores with increasing age. Rather no relationship was manifest between age and intellectual ability.

## References

1. Balinsky, B. An analysis of the mental factors of various age groups from nine to sixty. *Genet. Psychol. Monog.*, 1941, **23**, 191-234.
2. Brown, C. W., & Ghiselli, E. E. Age of semiskilled workers in relation to abilities and interests. *Person. Psychol.*, 1949, **2**, 497-511.
3. Ghiselli, E. E. The forced-choice technique in self-description. *Person. Psychol.*, 1954, **7**, 201-208.
4. Jones, H. E., & Conrad, H. S. The growth and decline of intelligence: A study of a homogeneous group between the ages of 10 and 60. *Genet. Psychol. Monog.*, 1933, **13**, 223-298.
5. Jones, H. E., & Kaplan, O. J. Psychological aspects of mental disorders in later life. *In* Kaplan, O. J., *Mental Disorders in Later Life.* Stanford: Stanford Univ. Press, 1945.

6. KUHLEN, R. G. Social change: A neglected factor in psychological studies of the life span. *Sch. & Soc.*, 1940, **52**, 14-16.
7. LORGE, I. The influence of the test upon the nature of mental decline as a function of age. *J. Educ. Psychol.*, 1936, **27**, 100-110.
8. MILES, C. C., & MILES, W. R. The correlation of intelligence scores and chronological age from early to late maturity. *Amer. J. Psychol.*, 1932, **44**, 44-79.
9. OWENS, W. A. Age and mental abilities: A longitudinal study. *Genet. Psychol. Monog.*, 1953, **48**, 3-54.
10. OWENS, W. A. The retest consistency of Army Alpha after 30 years. *J. Appl. Psychol.*, 1954, **38**, 154.
11. RAVEN, J. C. The comparative assessment of intellectual ability. *Brit. J. Psychol.*, 1948, **39**, 12-19.
12. VINCENT, D. F. The linear relationship between age and score of adults in intelligence tests. *Occup. Psychol., Lond.*, 1952, **26**, 243-249.
13. WECHSLER, D. The Measurement of Adult Intelligence. Baltimore: Williams and Wilkins, 1944.

*Department of Psychology*
*University of California*
*Berkeley 4, California*

*The Journal of Genetic Psychology*, 1957, **90**, 143-150.

# WORD VARIETY AS A MEASURE OF BILINGUALISM IN PRESCHOOL CHILDREN*

*Department of Psychology, University of Hawaii*

MADORAH E. SMITH

## A. PURPOSE

The purpose of this study is to determine if there is any difference in the number of different words per 100 as used by children from bilingual homes as compared with those from monolingual homes: and to try out the use of the number of different words in both languages used as a measure of the relative proficiency in each language.

## B. SUBJECTS

For this study use was made of 199 records of the conversations of two- to six-year-old monolingual children, most of which had been collected for previous studies (2, 3) and 101 records of bilingual children of the same age of Chinese ancestry which had been taken for another study (4). More records than those used had been recorded but all those less than 100 words in length were discarded and also from the bilingual records those which did not contain either 100 words in one of the two languages or 50 words in each. Most of these shorter samples thus discarded were of the younger or duller children so that the remaining records that were used are probably of children somewhat above the average.

## C. PROCEDURE

In a previous study (5) it had been found that *DW* or the number of different words per 200 or 500 running words had a low correlation with age for preschool children. Even 200-word samples were not available for our bilingual subjects for few of them used so many words in the 50-sentence samples recorded. Therefore the 100-word *DW* was used. Language ratings (*LR*) were given each bilingual child according to the relative amount of each language he heard. *LR* was calculated from data as to the languages spoken or preferred by each person older than himself in his home with a double weight given to that one spoken at kindergarten if he was an attendant there. The ratings ran from 1 to 7, practically all Chinese

---

*Received in the Editorial Office on January 3, 1956.

to nearly all English. Since the child's playmates could not be rated, but almost all our subjects preferred to speak English while playing with other children, the midpoint of *LR* is probably 3½ rather than 4.

## D. Results

The mean *DW* for monolingual two-year-olds as shown in Table 1 is 47.1, for five-year-olds 60.2; a highly significant difference. The difference in the means between succeeding age levels is not always significant but that from two to four and from three to five are both significant at the 1 per cent level of confidence. Chotlos (1) found a mean of 65.6 in the written work of school children averaging 129.8 months in age.

For the bilingual children two *DW*'s were calculated: the *ECDW* based on 100 running words whether they were English or Chinese and the *EC-D* with the same base but with words of duplicate meaning subtracted. Only for the former were $\sigma$'s of the mean calculated. Again the differences in *DW* for adjoining age levels are not significant, but the gain from three to five years is significant at the 1 per cent level of confidence.

At each age level, the monolingual children excel the bilingual in word variety, but the numbers at each level are too few for the differences to be significant. The difference at two years is the smallest; probably at this age the bilinguals are rather a select group since in discarding the shorter records most of the children whose fathers were of the lower occupational levels were lost. The next smallest difference is at five years, at which age the bilingual children average 1½ months, the older and 21 of the 36 children had started kindergarten, and only four of the 36 did not show a distinct preference as to choice of language as seen by the very few words of their non-preferred language spoken while under observation. Since only at four years did the two groups average the same age, children of the monolingual group were selected who matched one of the bilingual group as to sex and within two months of age so that the mean ages of the two groups were identical. It was impossible to find a match for one of the oldest children so that the number of pairs is 100, 59 of whom are boys and 41 girls. The mean *DW* of the monolinguals is 56.49 and of the bilinguals 53.72; a difference of 2.77 which is significant at the 2 per cent level of confidence. *DW* therefore shows, as does sentence-length and error index (4), that the bilingual preschool child is under a handicap that probably lessens as he grows older. The mean *EC-D* is 51.9, which is 1.6 lower than the *ECDW* which suggests that the bilingual child has an even smaller number of concepts that he can name.

The bilingual children were then separated into three groups according to language ratings. The children who heard the most English, over 60 per cent, with *LR*'s of 5, 6, and 7 were put in the first group, those with *LR*'s of 3 or 4 who heard approximately equal amounts of the two languages in the middle group, and those with *LR*'s of 1 or 2 who heard less than one-third English in the third group. At three and four years the children hearing most nearly equal amounts of English and Chinese have the lowest *DW* of the three groups. In this middle group there is but one two-year-old and when nine of the five-year-olds of this group are matched by age, sex, and kindergarten attendance with some from the highest *LR* group, it is again the middle group which has the lowest *DW*. Again it was not possible to match the oldest child. It would appear that for children under school age, it is more advantageous if there is a distinct preference in his home given to one of the two languages heard so that he does not hear them in nearly equal amounts.

Next the bilingual children were compared as to the amount of English used while under observation. By far the greater majority of them preferred to speak English while at play even if Chinese was more often spoken in their home. Only five used 20 per cent or less English, 24 from 27 to 70 per cent, and 72 at least 76 per cent English. Since the numbers in each group varied so much in number and mean age, matching was again resorted to and the 24 children speaking an appreciable amount of both languages were paired with others using over 90 per cent English while under observation. When matched the mean ages of each of the paired groups is 45.1 months, which closely agrees with the 45.0 month average of the five children who used mainly Chinese. On *ECDW* and *EC-D*, the last-named group ranks first by 4.0 and 2.3 respectively. Group *C*, the one using the most English, ranks second but their *ECDW* exceeds Group *B*'s by a negligible amount although their *EC-D* is 4.0 words the higher. As we wished to compare the number of different words in both languages, an *EDW* based on the number of different English words found in 50 English running words and a *CDW* calculated in the same manner for the Chinese words was calculated wherever possible for each child. Again Group *B* is at a disadvantage. Its *CDW* is 4.7 words lower than that of Group *A*'s and its *EDW* 8.4 lower than Group *C*'s. This suggests that the little child who is not trying too hard to use two languages has the richer vocabulary.

Breaking the matched pairs into three sets according to the amount of English used by the ones in Group *B*, the advantage of Group *C* over Group *B* is found to be the greatest for the members matching Set 2 who use a mean of 47.3 per cent English. Their *ECDW*, *EC-D*, and *EDW* exceed the corresponding scores of their pairs in Group *B* by 3.2, 4.6, and 1.2 words respectively. Of the pairs in Set 1, who use from 27 to 37 per cent English, Group *C* also has higher scores on all three measures but the difference is not so great. But in Set 3 where Group *B* members use an average of 66.8 per cent English, it is only on *EC-D* that Group *A* retains the advantage. However, if the average of Group *B*'s *CDW* and *EDW* is compared with *A*'s *EDW*, Group *A* has the higher score.

Only for Group *B* is it possible to compare *CDW* and *EDW* to try if these measures are indicative of the child's relative proficiency in the two languages. A correlation of .67±.07 was found to exist between the percentage of English spoken while under observation and the amount by which *CDW* exceeded *EDW*. The *CDW* of the children using the least English and whose *LR* is 22 is 4.7 higher than their *EDW*. Those using nearly equal amounts of the two languages have a mean *CDW* 4.3 higher than their *EDW*, but the *CDW* of those using 62 per cent to 70 per cent English is 3.3 lower. The difference between the *DW*'s for the four with the lowest *LR*'s is only one word, but for the other half of this set the *EDW* exceeds the *CDW* by 5.6. Since not only which language a child happens to use during a short period of observation but also the one which he hears most often are related to his degree of proficiency in the two languages it appears that the difference between the number of different words in equal sized samples of running words in the two languages is also a measure of his proficiency in each.

## E. SUMMARY AND CONCLUSIONS

1. The number of different words per 100 running words used by preschool children when at play with other children, 199 of whom were monolingual, was compared with 101 such records made by bilingual children of Chinese ancestry in the city of Honolulu. At each of four age levels from two to five years, the monolingual children used a greater variety of words. When 100 pairs were matched by age and sex, the monolingual children excelled the bilinguals by 2.77 words at a 2 per cent level of confidence.

2. When the bilingual children were separated into three groups according to the amount of each language heard, the children hearing more nearly equal amounts of each, tended to score lower on *DW* than those hearing either primarily English or mostly Chinese until after five years old when most of them were attending kindergarten and much of the languages they heard came from different sources, e.g., Chinese at home, English at kindergarten.

3. Most of the bilingual children used more English, the language of

Hawaii, than Chinese. When they were separated into three groups according to the percentage of English spoken while under observation no consistent difference in *DW* was found except that the few using mostly Chinese were consistently superior to their age mates. Children using at least 50 words in each language and from 27 to 70 per cent English, Group *B* were matched by age, sex, and kindergarten attendance and paired with children using 90 per cent or more English Group *C*. Then those children using mostly English or mainly Chinese scored higher on four types of *DW* than the ones using more nearly equal amounts of the two languages. When *EDW*'s were calculated, viz., *DW* based on 50 English running words, those using nearly all English scored higher. However when the pairs were broken down into three sets according to the percentage of English used by those in Group *B* the superiority of those in Group *A* was found to be due primarily to the children in the lower sets since the children in the set using on an average 67 per cent English did better on two measures than their mates in Group *C*.

4. Both *EDW* based on 50 English running words only and *CDW* on Chinese words only were calculated for all children using enough of both languages. The difference between *CDW* and *EDW* correlates .67±.07 with percentage of English used. Those children using less than 40 per cent English scored 4.7 words higher on *CDW* than on *EDW* and those using more than 60 per cent English 3.3 words lower on *CDW* than on *EDW*. The difference between the number of different words in equal-sized samples of words in different languages appears to be indicative of the relative proficiency in the languages used although it would be desirable to secure samples longer than 50 words.

## REFERENCES

1. CHOTLOS, J. W. Studies in language behavior: IV. A statistical and comparative analysis of individual written language samples. *Psychol. Monog.*, 1944, **56**, 77-111. (Whone No. 255.)
2. SMITH, M. E. An investigation of the development of the sentence and the extent of vocabulary in young children. *Univ. Iowa Stud.: Stud. Child Welfare*, 1926, **3**, No. 5. Pp. 92.
3. ———. A study of some factors influencing the development of the sentence in preschool children. *J. Genet. Psychol.*, 1935, **46**, 182-212.
4. ———. Some light on the problem of bilingualism as found in a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii. *Genet. Psychol. Monog.*, 1939, **21**, 119-284.
5. ———. Relations between word variety and mean letter-length of words with chronological and mental ages. (In press.)

*Department of Psychology*
*University of Hawaii*
*Honolulu, Hawaii*

*The Journal of Genetic Psychology*, 1957, **90**, 151-166.

# RELATIONSHIPS BETWEEN EARLY SOCIAL EXPERIENCE AND ADULT AGGRESSIVE BEHAVIOR IN INBRED MICE*

*Roscoe B. Jackson Memorial Laboratory*

JOHN A. KING[1,2]

## A. INTRODUCTION

A relationship between early social experience and adult aggressive behavior has been found in one inbred strain of mice (9). Male mice raised in isolation after weaning tend to be less aggressive adults than males raised in social groups. The effect of early social experience on adult aggressive behavior in mice may be used to test the critical period hypothesis (2, 12) and to discover the nature of the relationship between early experience and adult behavior.

According to the critical period hypothesis there are "specific stages in ontogeny during which certain types of behavior normally are shaped and molded for life" (2). The purpose of this study is to identify (*a*) the period or age when the social experience is most effective in producing adult aggression in mice, (*b*) the duration of the experience, and (*c*) the nature of the experience. Further objectives were to learn if the effects of the early experience (*d*) perseverate throughout life, (*e*) generalize into different adult activities, and (*f*) are genetically specific.

## B. SUBJECTS

There were 288 male mice of the C57BL/10 and BALB/c inbred strains used in this study. All mice were housed in 12-in. by 6-in. by 6-in. compartments of a wooden box containing two compartments. The tops of the compartments were covered by wire screen from which a food hopper and water bottle were suspended. The procedure describes the way the mice were reared and divided into groups.

---

*Received in the Editorial Office on January 9, 1956.

[1] This investigation was supported by research grant MH-123 from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

[2] The author wishes to acknowledge the assistance of Helen Connon in collecting the data and the staff of the Division of Behavior Studies, R. B. Jackson Memorial Laboratory for their suggestions during the course of the study: J. P. Scott, J. L. Fuller, R. Chambers.

## C. Apparatus

Mice were fought in boxes which were similar to the living cage except that the partition separating the two compartments was removable. During the test, the partition was removed, permitting the mice in each compartment to come together. No handling of the mice prior to the test was required.

## D. Procedure

### 1. *Constitution of Groups*

Groups I through VIII consist of male C57BL/10 mice. Group IX and X, which duplicate Groups I and V, consist of male BALB/c mice.

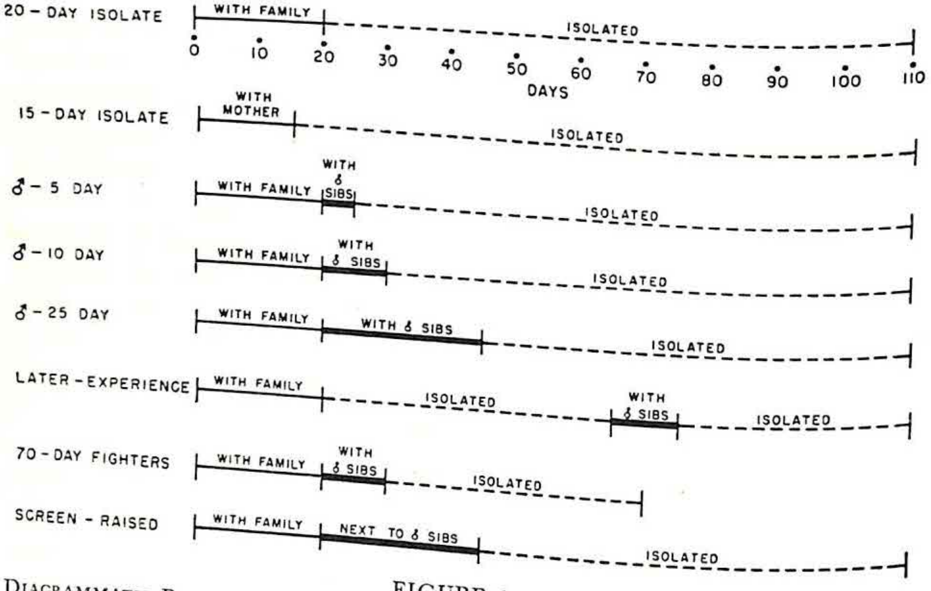

FIGURE 1

Diagrammatic Representation of the Treatments Given to Eight Groups of Mice Prior to the Aggression Test Given About 110 Days of Age

*a. Group I* (20-day-isolates). Forty-two males were weaned at 20 days of age and isolated from all other mice until placed in the testing situation at 100 to 120 days of age. This group is a combination of 24 mice from a previous study (9) and 18 additional mice which were given the same treatment. A graphic representation of the early experience of this group compared with seven other groups is presented in Figure 1.

*b. Group II* (15-day-isolates). Twenty-two males were raised singly (other litter mates were discarded) by their mothers until 15 days of age, at which time they were separated from their mother and kept in isolation until placed in the testing situation at 100 to 120 days of age. This treatment eliminated sibling or paternal associations before weaning and maternal influences at the earliest possible age. The 22 mice in this group are the survivors of an initial 60 mice which were weaned at 15 days.

*c. Group III* (♂-5-days). Twenty-four males were separated from their parents and all female siblings at 20 days of age and raised with only their male siblings until 25 days of age, at which time they were isolated until placed in the test situation at 100 to 120 days of age.

*d. Group IV* (♂-10-days). Twenty-eight males were treated like Group III mice except they remained with the male siblings from 20 to 30 days of age.

*e. Group V* (♂-25-days). Fifty-two males were treated as the preceding two groups, except they were left with their male siblings and father from 20 to 45 days of age. Data from a preceding study using 34 mice were added to the data from 18 mice in the present study. The addition of the fathers as a variable in this group made the young no more aggressive than those raised only with male siblings, as shown in the results.

*f. Group VI* (later-experience). Thirty-four C57BL/10 males were isolated from other mice at 20 days of age until they were 65 days old, when they were placed with their male siblings, which had also been isolated. They remained together for 10 days and were isolated again at 75 days of age until placed in the test situation at 100 to 120 days of age.

*g. Group VII* (70-day-fighters). Sixteen males were treated like Group IV mice, except that the interval of isolation between the social experience and the test situation was reduced to 40 days. This reduction in the duration of isolation necessitated fighting the mice at 70 days of age.

*h. Group VIII* (screen-raised). Twenty male C5BL/10 mice were weaned at 20 days of age and were placed alone in one compartment of a living cage with a double ¼ inch wire mesh screen instead of a solid wooden partition separating the isolates. The screen prevented physical contacts between the mice. After 25 days (45 days old) of this semi-isolated treatment, the mice were totally isolated until placed in the test situation at 100 to 110 days of age. This treatment corresponded to Group V mice except actual contact was prohibited.

*i. Group IX* (20-day-isolate BALB/c). Twenty-four male BALB/c mice were given the same treatment as Group I in the C57BL/10 mice.

*j. Group X* (♂-25-day BALB/c). Twenty-six male BALB/c mice were given the same treatment as Group V in the C57BL/10 mice.

### 2. *Procedure of Aggression Tests*

Two members of the same group were placed individually in each compartment of a fighting box three days prior to being tested. Pairs of combatants were formed from mice that had never lived together. The tests began by removing the partition between the two compartments and the mice could then come together. The frequency of social interactions (running away, nosing, smelling genitals, grooming, attacking) from the time the mice were exposed to each other until the first fight began was recorded. Fights were distinguished from attacks by the participation of both mice in a fight. The principal measure of aggressive behavior used in this study was the fighting response latency, which is the elapsed time from the moment when the partition was removed until the first vigorous fight began (4). The pair was allowed to fight approximately 5 seconds before the partition was lowered between them. Often the combatants had to be separated with forceps, although most battles were brief and the mice separated temporarily before returning to fight. Dominance was prevented by separating the mice after 5 seconds of fighting. A maximum of 5 minutes was allowed for each trial if the mice failed to fight. Each pair was given 10 trials at the rate of one each day. Since all mice in each group were not born on the same day, they were tested between 100 and 120 days of age with a median of approximately 110 days of age.

The response latency seemed the most satisfactory measure of aggression because it is reliable (a product moment correlation between even and odd trials gives $r = .99$, $N = 72$), it can be used without permitting dominance to form, it indicates the duration of a fight (3), and it is not affected by chance factors, such as a random bite which may terminate the fight by causing the injured mouse to submit.

The failure of the response latency data to meet the assumptions of normality required by parametric statistics necessitated the use of a nonparametric statistic. For this purpose the Mann-Whitney $U$ test (11) was applied.

### 3. *Procedure of Sexual Tests*

Ten mice from Groups I, II, IV, V, and VIII, were given sexual tests approximately three months after the aggression tests, during which time they were kept in isolation. The mice were tested by placing one estrus female into the living cage of each male. Females were brought into estrus by .01 mg. injections of estradiol (Progyron-B, Schering Co.) and tested for receptivity with a sexually active male until intromission was achieved.

The females were then placed with the male mice of each group for 10 minutes each day for 10 days. The frequency of intromissions are the only measure of sexuality activity used in this experiment because ejaculations were too infrequent and mountings do not represent as high a level of sexual activity as intromission.

## E. RESULTS

### 1. *Effect of Isolation*

The effect of isolation may reveal the influence early social relationships have upon behavior. Mice isolated at 20 days of age until tested for fighting at 110 days of age were inhibited in their aggressive behavior (9).

TABLE 1
STATISTICAL DIFFERENCES IN THE SECOND† TO TENTH TRIALS OF EIGHT GROUPS OF MICE RAISED WITH DIFFERENT EARLY SOCIAL EXPERIENCES

| Group | VI | III | II | IV | VII | VIII | V | Mean latency in min. |
|---|---|---|---|---|---|---|---|---|
| I | NS | NS | NS | *<br>*2 3* 4<br>*5 6 7*<br>*8 9 10* | *2 3 4*<br>*5 6 7*<br>*8 9 10* | *2 3 4*<br>*5 6 7*<br>*8 9 10* | *2 3 4*<br>*5 6 7*<br>*8 9 10* | 3.32 |
| VI | — | NS | NS | *2 3* 4<br>*5* 6 *7*<br>8 *9 10* | *2 3 4*<br>*5* 6 7<br>8 9 *10* | 2 3 *4*<br>*5 6 7*<br>8 *9 10* | *2 3 4*<br>*5 6 7*<br>*8 9 10* | 3.16 |
| III | | — | NS | 2<br>6<br>9 | 2 3 4<br>5<br>10 | 4<br>5 6<br>8 9 10 | *2 3 4*<br>*5 6* 7<br>*8 9* 10 | 3.03 |
| II | | | — | NS | 10 | 3<br>9 10 | 2 *3*<br>5 6 7<br>8 9 10 | 2.91 |
| IV | | | | — | NS | NS | 7 | 1.96 |
| VII | | | | | — | NS | NS | 1.72 |
| VIII | | | | | | — | NS | 1.73 |
| V | | | | | | | — | 1.19 |

†No significant differences in the first trial.
*Figures in italics indicate trials which reached the one per cent probability level, others indicate the five per cent level as determined by the Mann-Whitney *U* test.

When the aggressive behavior of mice isolated at 15 days of age (Group II) is compared with the behavior of mice isolated at 20 days (Group I), it can be seen from Table 1 that there is no significant difference between them. It was expected that earlier isolation would reduce aggressiveness even more. However, the mean response latency of Group II was less than that of Group I, which indicates a tendency for the earlier isolated mice to be even more aggressive than the later isolated mice. It is possible that this tendency of the Group II mice to be more aggressive is the result of a selection of the strongest and most aggressive mice, since about two-thirds of the initial group died during the treatment.

The inhibition of aggression by isolation, whether at 15 or 20 days of age, is apparent in the slower response latencies of Groups I and II than in Group V mice (Table 1).

### 2. *Effect of Duration of Social Experiences*

The duration or quantity of an early experience which an organism must receive before it is permanently engraved upon its behavioral pattern may be limited to one traumatic experience, may extend throughout certain developmental periods, or may recur at widely separated intervals throughout its development. The purpose of treating Groups III, IV, V, and VI with varying lengths of early social contacts was to ascertain the minimum number of days necessary to result in low fighting response latencies in adult mice.

The treatment of Group III differed from Group I by delaying isolation five days until the mice were 25 days old and by restricting their social contacts between 20 and 25 days of age to their male siblings. This difference in treatment was not sufficient to distinguish Group III mice ($\bar{x} = 3.03$ min.) from Group I mice ($\bar{x} = 3.32$ min.), although it did tend to reduce Group III fighting response latencies somewhat as compared to Group I (Figure 2).

When mice were given 10 days of social contacts with their male siblings (Group IV, $\bar{x} = 1.96$ min.) after weaning at 20 days and then were isolated, they were significantly (Table 1) faster fighters than mice isolated at 20 days of age (Group I). On the other hand, their response latencies did not differ significantly from Group V ($\bar{x} = 1.19$ min.) mice, which experienced 25 days of sibling and paternal contacts after weaning. These results indicate that social contacts for a period of five days immediately after weaning was not long enough to make the mice more aggressive than the completely isolated mice (Group I), but that a period of 10 days of social contacts produced about the maximum amount of aggressive behavior. Longer periods of social contacts were not necessary.

### 3. *Effects of Age of Social Experience*

If 10 days of social interaction immediately after weaning is sufficient to produce low fighting response latencies in adult mice, we may ask if the same amount of social interaction during adulthood has the same effect.

According to the critical period hypothesis, there are certain "periods in development when the same experience may have a more profound effect than at other times" (12). On the basis of this hypothesis, the 10 days of social exposure after weaning should produce different effects than 10 days of the same exposure after the mice have become sexually mature. Only one age, other than the 10 days after weaning, was tested, with Group VI mice. The age selected was 65 days, when most development has ceased, although the mice do continue to grow in size.

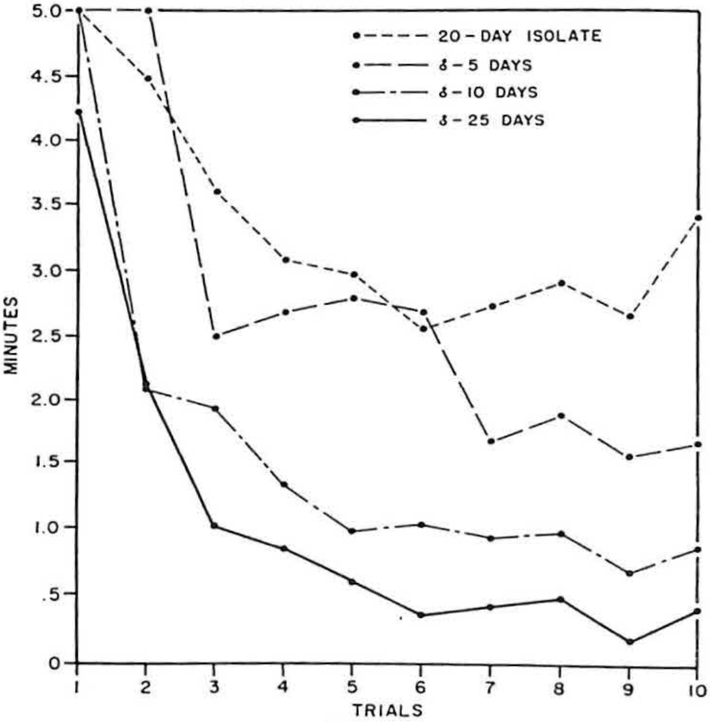

FIGURE 2

Median Curves of the Fighting Response Latencies of Groups I (20-Day-Isolate), III (♂-5-Day), IV (♂-10-Day), and V (♂-25-Day) of C57BL/10 Mice

Group VI mice had response latencies closer to Group I mice than any other group and had significantly slower latencies than Group IV mice in nine of the 10 trials as shown in Table 1. Since the experiences of the VI mice were similar to those of Group I mice immediately after weaning when both groups were isolated, we may conclude that this period is more critical in the development of aggressive behavior than at maturity when the mice experienced social contacts similar to those of Group IV mice.

### 4. *Effect of Duration of Isolation*

In order to keep the age at testing as constant as possible, it was necessary to vary the length of the period of isolation in the different groups. It is possible that this uncontrolled variable was contributing to some of the differences among the groups. One method of testing the effects of the duration of the isolation period is to change it radically and fight the mice at a different age. If the duration of isolation is contributing to the effects produced on adult behavior, two groups of mice treated alike, except for a different duration of isolation, should behave differently as adults.

Group VII mice, which were given the same social exposure as Group IV mice, but tested after a 40-day isolation period instead of an 80-day isolation period, did not differ from Group IV in any of the trials (Table 1). Within the limits tested, this variation in the isolation period cannot be considered to influence the adult aggressive behavior.

### 5. *Nature of the Social Experience*

In the preceding experiments, no attempts were made to understand the nature or the quality of the social experience obtained by the mice given periods of social contacts. Since no special training was given and no direct observations were made of the mice during the period they were living together, the results can only offer circumstantial evidence as to what actually occurred in the social groups. A preceding paper (9) suggested two alternative types of experience which may be influential in determining the adult behavior. One type of experience may be that of competition for food, water, or space, which acts as a type of latent learning for adult aggression. Williams and Scott have observed that conflict behavior over food does occur often between 25 and 40 days of age (15). The other type of experience may simply be one of familiarization with other mice, so when they come together in a fighting situation they are not inhibited by the strangeness of another mouse.

In order to ascertain which of these experiences is affecting the adult behavior, a group of animals was given the opportunity to become acquainted with other mice through a screen for 25 days after weaning, but not allowed to compete. If a competitive situation was necessary for the mice to learn aggressiveness, one would expect that the Group VIII mice would be slow to fight. On the other hand, if the experience needed was only some familiarity with another mouse, one would expect the Group VIII mice to fight rapidly when tested.

A comparison between Group VIII and Groups IV and V as shown in

Table 1 and the respective curves of the three groups in Figures 2 and 3 indicates that there was no significant difference among them. The unrestricted contacts among the siblings of Groups IV and V mice during the period of social living failed to distinguish them from Group VIII mice, which could only perceive each other through a screen during the same period. This failure of Group VIII mice to differ from the socially raised mice indicates that competition or any learning of aggression is not essential for developing a mouse eager to fight at maturity. Apparently the perceptual cues obtained through two wire screens was sufficient for establishing the pattern of behavior pursued at adulthood. These results suggest that the kind of experience obtained by the mice is that of familiarization with another mouse. It might be expressed in terms of an investigative drive. Once the investigative drive of a mouse is satisfied, it no longer serves to inhibit the tendency to fight.

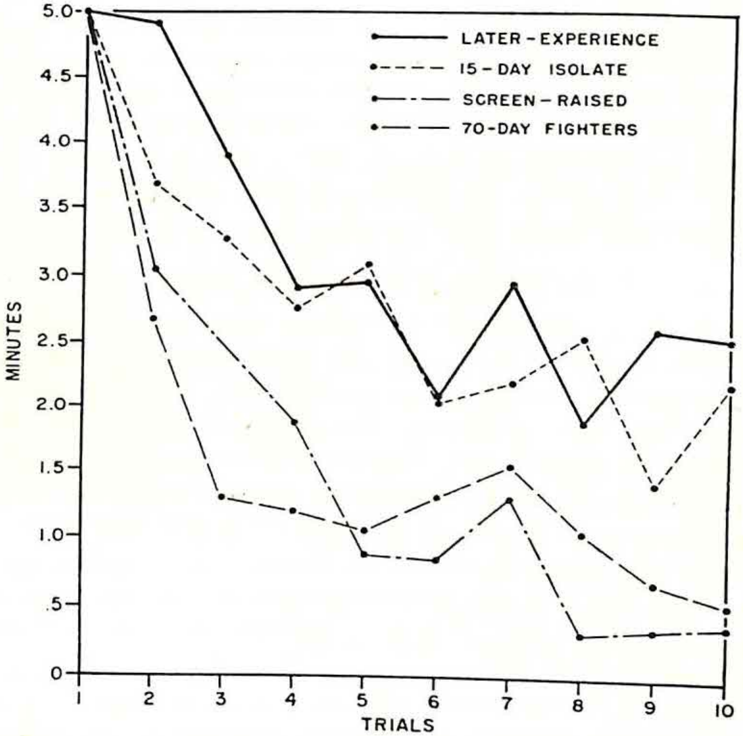

FIGURE 3
MEDIAN CURVES OF THE FIGHTING RESPONSE LATENCIES OF GROUPS VI (LATER-EXPERIENCE), II (15-DAY-ISOLATE), VIII (SCREEN-RAISED), AND VII (70-DAY-FIGHTER) OF C57BL/10 MICE

### 6. *Perseverance of the Effects of Early Experience*

Since all of the groups were tested for aggression only once in their lifetime, it seemed desirable to learn if the effects would persevere until a later age. In order to test how long the effects of the early experience persists after their first encounter and an intervening period of isolation, the Group III mice were retested at 200 days of age. The mice were isolated for the 90-day interval between the tests.

In all of the nine trials used for comparison, the mice fought significantly faster ($P < .01$ on Trials 3, 4, 6, and 8; $P < .05$ on Trials 2, 5, 7, 9, and 10) in the retest at 200 days of age than on the first test at 110 days of age. This significant change in the behavior of the mice after a period of three months indicates that the effects did not persist in the mice as they became older. A product moment correlation of the mean response latencies for each pair of mice in both the 110-day and 200-day tests showed a positive correlation ($r = .73$; $P < .01$) between the two tests. The mice which had short latencies at 110 days tended to have shorter latencies at 200 days. It is likely that the experiences of the first test influenced the behavior of the mice at a later age. Age differences are probably not alone responsible for the increased aggression because Group VII mice are just as aggressive at 70 days as Groups IV and V at 110 days of age.

### 7. *Generalization of Effects*

If the early social environment of the mice can affect their adult aggressive behavior, will the same experiences generalize into other types of bevior? From the work of Stone (13), Beach (1), and Valenstein (14), we know that different social environments of developing rats and guinea pigs can affect the adult patterns of sexual behavior. It seems likely, therefore, that the sexual behavior of mice raised with different social experiences after weaning may also be affected. The further question may be asked: if the social experiences in infancy tend to make the mouse aggressive, will he be more or less likely to exhibit high sexual activity?

In order to learn if the effects of early social experience generalize into types of behavior other than aggression, 10 mice of the following groups were given sexual tests approximately three months after the fighting tests: Group I, II, IV, V, and VIII. During the three-month period between the fighting and sexual tests, the mice were kept in isolation. This long interval between tests would tend to eliminate the carry-over of one test into the other. The use of the same mice for both tests permitted a correlation between the sexual and aggressive behavior of each individual to be made.

Group VIII mice had significantly fewer ($P < .005$) intromissions than any of the four other groups when the total number of intromissions over 10 trials were tested by the Mann-Whitney $U$ test. The Group II mice had fewer intromissions ($P < .05$) than Group V mice. These data demonstrate that Group VIII mice were less active sexually than the other groups and Group II mice were intermediate in their sexual activity. These differences cannot be attributed exclusively to early experiences because the mice also had the later experiences of the fighting tests. The experiences which produced an aggressive mouse did not necessarily enhance or inhibit its sexual behavior. For example, Group VIII mice, which were as aggressive as Group V mice, were less active sexually than Group V mice. The experiences which affect adult aggressive behavior apparently are specific and do not generalize into sexual behavior (10).

### 8. *Genetic Effects*

The genetic characteristics of the mice in Groups I through VIII were held constant by the use of a homozygous strain of inbred mice. Any differences in their behavior are consequently the result of their different environmental treatments. Before concluding that these treatments would affect animals of different genetic constitution in the same manner, it was necessary to try the same treatments on a different inbred strain. For this purpose, mice known to behave differently in aggressive situations (5), BALB/c, were given the same treatments as the two most divergent groups among the C57BL/10 mice. Group IX in the BALB/c mice corresponds in treatment to Group I of the C57BL/10 mice and Group X BALB/c mice corresponds to Group V C57BL/10 mice.

The slow fighting response latencies of the isolated Group I failed to appear in the Group IX mice, which were given the same early experiences as Group I. On the other hand, Groups V and X were similar in the adult aggressive behavior. The two BALB/c groups showed no significant differences in any of the trials (Figure 4). Both Groups IX and X tended to be very aggressive, reaching an asymptote in their response latency curves below that of any group in the C57BL/10 strain. The faster fighting exhibited by Group IX than by Group I mice may be attributed to genetic differences, since both groups were treated alike.

## F. DISCUSSION

The purpose of this study was to investigate the relationship between early social experience and adult aggressive behavior in mice in order to test the critical period hypothesis.

Since the age at which the mouse undergoes the experience is crucial to the hypothesis, it was necessary to learn if the same social exposure at different ages of the mouse have the same effect on its adult behavior. The results indicated that the age of the mouse was important because 10 days of social contact with their siblings following weaning had more effect on their adult aggressive behavior than 10 days of social contact with their siblings after the mice had become sexually mature. Although the same opportunities for social contacts were provided the mice at the different ages, it is impossible to conclude that the nature of the social contacts was similar. Mice 20 to 30 days old have different types of social interactions than mice 65 to 75 days old. Social exposure during maturity probably induces the siblings to fight and establish a dominance hierarchy. Further fighting among the defeated mice in the test situation may be inhibited. These defeated mice, which are not aggressive, may account for the long latencies in Group VI. The fact

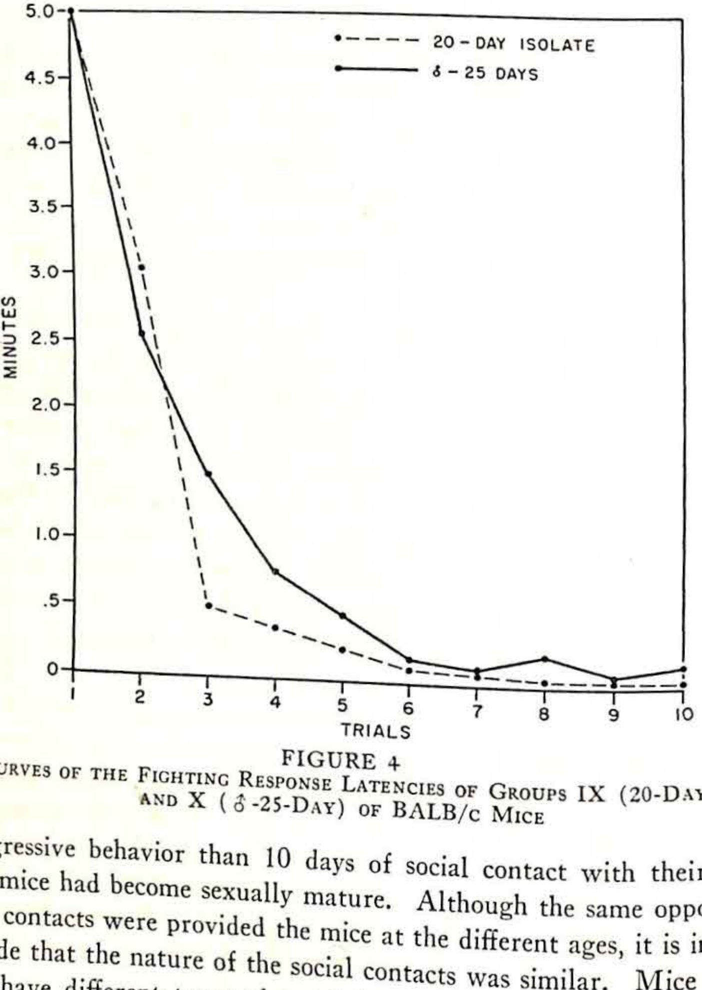

FIGURE 4

Median Curves of the Fighting Response Latencies of Groups IX (20-Day-Isolate) and X (♂-25-Day) of BALB/c Mice that mammals are capable of exhibiting different behavioral patterns throughout their maturation does not invalidate the critical period hypothesis.

The quantity and quality of the experience also plays a rôle in determining the effect of the early experience. One purpose of this study was to learn the minimum number of days of early social contacts necessary to establish a high level of aggression in adult mice. From the three lengths of social exposure examined in this study, namely, five, 10, and 25 days following weaning, we may conclude that longer periods have greater effects with diminishing returns. Mice exposed to siblings the first five days after weaning fight slower than those exposed the first 10 days, and mice exposed 10 days are slightly, but not significantly, slower to fight than mice with 25 days of exposure. Since the differences in the fighting latencies are greatest between five and 10 days, we may conclude that 10 days of social contacts immediately after weaning is the minimum length of time needed to produce adult patterns of aggressive behavior.

The quality of the early experience was not directly observed in the present study in an effort to allow the social associations to be as natural as possible. In an experiment of Fredericson (5), one group of mice trained to compete for food between 29 and 36 days of age and another group for just one day at 33 days of age were more competitive as adults than another group without the training. Kagan and Beach (7) observed that the type of behavior of young male rats toward receptive females influenced the adult pattern of male sexual behavior. Kahn (8) subjected groups of young mice to 30 days of severe defeats by another mouse starting at 21, 35, and 60 days of age. He found the greatest differences in adult aggressive responses between mice which had the experience at 21 and 60 days of age. In the present study the nature of the social experiences of young mice living together was analyzed by restricting their contacts with a screen partition. This prohibition of overt fighting and competition for food, water, and space did not result in low fighting latencies in the adults. The contacts made through the screen were sufficient to produce maximum fighting responses. It appears that the slow fighting response latency of isolated mice results from an inhibiting mechanism. In a preceding paper (9), social investigation was suggested as the mechanism which inhibited fighting. Before mice fight, they must first investigate each other. The duration of this investigatory period in the test is prolonged among the isolated mice because they have not seen another mouse since weaning at 20 days. In contrast, those mice which have had the opportunity to investigate other mice through a wire screen for only 10 days following weaning have acquired sufficient experience with each other to react aggressively in a fighting situation.

Although the effects of early social experience may perseverate into adulthood when the mice are kept in isolation, the results from retesting Group III at 200 days of age suggest that intervening social experiences may eliminate the early effects. Since the acquisition of social experiences is a continuous process in the normal development of most mammals, the effects of the early experience may be altered by subsequent experiences.

Early experiences may affect adult aggressive and sexual behavior in mice (5, 8), rats (7), and guinea pigs (14), but there is no evidence that the effects of the early experience generalize from one type of adult behavior to another type. It has not been proven possible to predict the effect of an early experience on adult sexual behavior from the effects produced on adult aggressive behavior (10). An early experience which induces aggressiveness in adult mice may either inhibit or enhance their adult sexual behavior.

Genetic factors appear to be important in assessing the effect of early experience on adult behavior. A mouse of one genotype may be affected by an early experience differently than a mouse of another genotype. These results illustrate the need for caution in accepting the effect of early experiences as a general principle. In any heterogeneous society of mammals, some individuals may be more predisposed to modification by a given early experience than are others.

Although the results of this study confirm the critical period hypothesis in general, they also indicate that there are several qualifications in its application to mice. (*a*) The effects of the experiences gained during the critical period may be altered, reduced, or eliminated by subsequent experiences; (*b*) they may be specific for each type of adult behavior; and (*c*) there may be different effects on individuals of unlike genotypes. These qualifications suggest a modification of the critical period hypothesis previously quoted (2). There are specific stages in ontogeny during which certain capacities for behavior appear; the manner in which these capacities are utilized determines subsequent behavior; if these capacities are not utilized they tend to be inactivated.

## G. SUMMARY

A total 288 mice of C57BL/10 and BALB/c strains were given different social experiences or the same experience at different ages in order to learn the effect of these experiences on adult aggressive behavior. Adult aggression was measured by the fighting response latency of mice initially exposed to each other. The following statements pertain only to the strains of mice used in this study.

1. Male mice with 10 or 25 days of social associations immediately after weaning are faster fighters than males with five days or without postweaning social associations.

2. Male mice which have social contacts immediately after weaning are faster fighters when adults than mice which have social contacts during maturity.

3. Male mice separated by a screen mesh from their siblings for a period of 25 days after weaning show no difference in aggression from those raised in contact with each other for the same period.

4. A second aggression test repeated 90 days after the first test indicates that male mice are faster fighters in the second test. The effect of the early social associations did not perseverate.

5. Male mice which were given both sexual and aggressive tests failed to demonstrate that their sexual and aggressive behavior were similarly affected by their early experiences.

6. Male C57BL/10 mice isolated at 20 days of age were less aggressive adults than male BALB/c mice isolated at the same age.

## References

1. Beach, F. A. Comparison of copulatory behavior of male rats raised in isolation, cohabitation, and segregation. *J. Genet. Psychol.*, 1942, **60**, 121-136.
2. Beach, F. A., & Jaynes, J. Effects of early experience upon the behavior of animals. *Psychol. Bull.*, 1954, **51**, 239-263.
3. Collias, N. Statistical analysis of factors which make for success in initial encounters between hens. *Amer. Nat.*, 1943, **77**, 519-538.
4. Fredericson, E. Response latency and habit strength in relationship to spontaneous fighting in C57 black mice. *Anat. Rec.*, 1949, **105**, 29 (abstract).
5. ———. Competition: The effects of infantile experience upon adult behavior. *J. Abn. & Soc. Psychol.*, 1951, **46**, 406-409.
6. Ginsburg, B., & Allee, W. C. Some effects of conditioning on social dominance and subordination in inbred strains of mice. *Physiol. Zoöl.*, 1942, **15**, 485-506.
7. Kagan, J., & Beach, F. A. Effects of early experience on mating behavior in male rats. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 204-208.
8. Kahn, M. W. The effect of severe defeat at various age levels on the aggressive behavior of mice. *J. Genet. Psychol.*, 1951, **79**, 117-130.
9. King, J. A., & Gurney, N. L. Effect of early social experience on adult aggressive behavior in C57BL/10 mice. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 326-330.
10. King, J. A. Sexual behavior of C57BL/10 mice and its relation to early social experience. *J. Genet. Psychol.*, 1956, **88**, 223-229.
11. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stockastically larger than the other. *Ann. Math. Statist.*, 1947, **18**, 50-60.
12. Scott, J. P., & Marston, M. Critical periods affecting the development of normal and mal-adjustive social behavior of puppies. *J. Genet. Psychol.*, 1950, **77**, 25-60.

13. STONE, C. P. The congenital sexual behavior of the young male albino rat. *J. Comp. Psychol.*, 1922, **2**, 95-153.
14. VALENSTEIN, E. S. The rôle of learning and androgen in the organization and display of sexual behavior in the male guinea pig. *Amer. Psychol.*, 1954, **9**, 486 (Abstract).
15. WILLIAMS, E., & SCOTT, J. P. The development of social behavior patterns in the mouse, in relation to natural periods. *Behaviour*, 1953, **6**, 35-64.

*R. B. Jackson Memorial Laboratory*
*P. O. Box 847*
*Bar Harbor, Maine*

*The Journal of Genetic Psychology*, 1957, **90**, 167-202.

# SOME HYPOTHESES ON PERCEPTUAL AND LEARNING PROCESSES WITH THEIR APPLICATIONS TO THE PROCESS OF READING: A PRELIMINARY NOTE*

*Reading Improvement Program, Tulane University*

HELENE POWNER MANN

## A. THE PROJECT

This is a report of some hypotheses with respect to perception and to learning generally. The hypotheses grew in part out of clinical and teaching experiences first, and more recently out of a program for the improvement of reading which was begun at Tulane University in 1953. Only a beginning has been made towards evaluating them specifically in terms of experimentation and in some instances even this has not been possible. The present report is but a preliminary note of hypotheses, the evaluation of which is obviously dependent upon a wide and long-range series of projects. Reading as a culturally important area of perception and learning provides a useful medium for the study of these subjects.

The plan of this paper is as follows: It will include a report of a reading improvement program together with some of the measurable and non-measurable results of the first year and observations made upon the students. It was the interpretation of these which led to some of the hypotheses which follow. It will also include some of the reading program plans which were and are desirable but which could not be implemented this year. Finally, the conclusions are stated together with some hypotheses. Of the latter, some are believed to be only incidentally applicable to the reading process. It is hoped that our experiences may be of use to other universities and colleges which may be contemplating a reading improvement program and to those who are open-mindedly seeking answers to related problems.

## B. THE PROBLEM

The problem was: To determine measurable and non-measurable results in a university reading program consisting of 12 50-minute periods supplemented by individual conferences within the time limits of the psychologist and students. Twelve was a somewhat arbitrary number. It was expedient at that time to divide the semester into as many series of classes as possible

---

*Received in the Editorial Office on January 10, 1956.

in order to assist a maximal number of students in classes small enough for individual attention and group interaction.

## C. The Program Plan

The program for the improvement of reading is designed essentially for freshman students who are deficient in reading but not confined to them. It is to be considered a part of a larger university plan to assist students to improve academic work, to reduce the number of failures, to increase the opportunities for wider experiences in college activities without diminishing academic success, to increase the time for leisure of the student's own choosing and to enhance leisure by making it possible to enjoy participation in a sphere of ideas and books as well as of people and things. These objectives are facilitated by helping the students to acquire reading habits which enable them to read with understanding at rates which are commensurate with individual capacities, a condition which does not necessarily exist of itself.

The essentials of the reading program plan were as follows: (*a*) Individual diagnosis of reading problems followed by appropriate remedial instruction given individually or in small groups according to the nature of the problem.

(*b*). The program to be voluntary and confined to students of the university for the present.

(*c*). The program to concentrate its efforts this year upon freshmen and urgent cases above the freshman level.

(*d*). A standardized test of reading to be administered to the entire freshman class for survey and screening purposes.

(*e*). The results of the standardized tests to be conveyed to freshmen through a mimeographed and explanatory letter with recommendation to those below a specified level that they consult with the director of the program.

(*f*). The program to be integrated with other university facilities contributing to student welfare, particularly those allied to physical and mental health.

(*g*). A direct line of coöperation to be established with student health services and the Medical School with a view to obtaining physical examinations, eye examinations, and hearing tests for students whose reading appears to be influenced by physical factors. Ideally, this would be done in all instances.

## D. THE PROGRAM AS CARRIED OUT

In order to screen freshman students with reading disadvantages, a survey of the freshman class was made using the Coöperative Reading Test, *C 2*, Form *Z*. Those students above the freshman year who were interested in taking the test were invited to do so. Eight hundred thirty-six students, of which 764 were freshmen took the test. The results were typed in, on mimeographed forms which the students were given together with a personal explanation when they inquired for their results. An announcement of their availability was made through the college paper and through freshman assemblies. Students whose scores were below the fifteenth percentile were encouraged to enroll in a reading class. The results of those whose scores placed them in the lowest five percentiles and later the lowest 14 percentiles were sent to the deans of the colleges with the recommendation that the students be encouraged to enroll in a reading group. Eventually the complete results and a statistical summary were sent to the deans.

The policy was to accept any student who came for assistance in reading. One hundred fifty-three students were enrolled in classes or given equivalent consultation. Eleven classes in reading improvement were held in series of 12 meetings each. Conferences preceded and supplemented class work and when the latter was not feasible were substituted for it.

In addition to the reading classes, five classes were held the last two weeks of the first semester for purposes of instruction in "How to Study for Final Examinations," a small group of foreign students was given instruction in methods of acquiring a reading and speaking vocabulary, and two students were given 20 hours each of individual reading instruction.

The classes in "How to Study for Final Examinations" given the first semester were the source of a number of enrollments in the classes in reading improvement which began the second semester. The students were informed that last minute study was superficial and that they could not expect to retain for long material learned at this time. However, they were given practical suggestions for study and reading which it was hoped would be carried over into future work. The principal reason for giving these courses was the newness of the reading program and the importance of letting the students know about it through experience with it.

The class for foreign students was given at the request of a member of the English Department and because of the writer's experience in teaching foreign languages. It was definitely not a course in the improvement of reading excepting as the principles of vocabulary-building with practice are basic to good reading. These students were taught how to learn English vocabulary most efficiently through application of visual, auditory, and kinaesthetic principles. Since the foreign students presented a special problem, no attempt was made to continue beyond these fundamentals, and the students were referred back to the remedial English classes.

## E. The Reading Classes

The reading classes comprise the group upon which the statistical data are based. They met 12 times each and were supplemented by individual conferences within the time limits of the students and the psychologist. Some of the groups met three times weekly, and some twice, so that the course lasted a month or six weeks depending upon the group in which the student enrolled. The reading rooms were open from nine until five o'clock for use of those who were able to make time to do extra practice in their already full schedules.

Each student was given the Iowa Silent Reading Test first as a part of the diagnostic procedure and again at the completion of the course in alternate and equivalent form as a measure of progress. This test includes nine areas of reading. While it is less of a depth test than the Coöperative Reading Test, it has some diagnostic advantages used in conjunction with the Coöperative Test. It would have been desirable to repeat both tests but a compromise had to be made in view of the fact that this was a service rather than a research program.

The coöperation between the reading program and the student counselling service made it possible to have individual tests of intelligence, additional reading tests, and vocational counselling in some cases. Vision was tested on the Keystone Visual Survey Telebinocular when indicated or requested and routinely when time allowed. The Speech and Hearing Center gave hearing tests in a sound-proof room by means of a puretone audiometer in some cases. Students were unfortunately not always sufficiently motivated to make the trip downtown to the Medical School but the urgent cases did. We were not successful in obtaining visual examinations through the Eye Clinic of the Medical School since the Director was away. Work was also not begun in coördinating the reading program with student health services.

The students were told at the beginning of the course that they would be treated as adults, that it was assumed that they did not require nor desire a mother-substitute to exert pressure upon them in college, but that if their own goals were not sufficiently strong or urgent in the direction of reading improvement this would be respected. They were encouraged to make use of the Guidance Center for vocational counselling, and a number of them did.

It was pointed out to them that each individual in the group differed from all of the others with respect to his reading and, that, therefore, they would each one be in a non-competitive position. They were to accept their reading problems as they were revealed through tests and observation to the psychologist and to themselves, and from these starting points would attempt to improve by means of the suggestions and insights received in class and in conference. It was further explained that, due to numbers, some initiative would be required of each with respect to asking for appointments for conference and for seeing that they got the most out of the course.

Since attendance was voluntary and was encouraged but not required, the students were free to withdraw at any time or to omit classes. The emphasis was not upon attendance but upon improvement of reading. During the second semester, when the reading classes opened, 51 students were in attendance at a first meeting when the Iowa Silent Reading Test was administered and for the alternate form given at the end of the course. Forty-eight others enrolled and came to at least three meetings or came intermittently for individual work as time allowed. Some dropped out altogether. Those who dropped out appeared to group themselves somewhat as follows: (*a*) Students whose results on the Iowa Silent Reading Test surpassed their expectations or their results on the survey test given in the fall. (*b*) Superior students who quickly observed the essentials to improved reading and found class procedure too slow. A number of these continued to come intermittently or for conference or both. (*c*) Students whose motivation appeared to be strong in the direction of other work but not reading. (*d*) Students whose motivation appeared to be stronger on a play level than on a work level. At least four students dropped out because of disappointment over the discovery that reading improvement required effort and could not be made solely by the use of mechanical devices which they had heard or read about.

## F. THE CONTENT OF THE READING COURSE

The content of the course was:

*Vocabulary building*, with exercises involving prefixes, suffixes and roots, and keeping a special section in their notebooks for an active vocabulary in each college subject in the form of check lists.

*Outline making*, with most of the students making two short outlines, to determine whether they had grasped the principles of outlining. In some cases the second outline was not necessary. They were encouraged to use outlines in studying their college texts and in note-taking.

*Graph making*, with each student making a graph of his results on the Coöperative Reading Comprehension Test and the Iowa Silent Reading Test. Those who took the alternate form of the Iowa Silent Reading Test at the end of the course superimposed their terminal results upon the graph of their initial results. In this way, they could note, within the limits of the test, where their strengths lay and where their efforts appeared to be most required. They kept the graphs as final reports and were encouraged to continue reading practice in the direction suggested by the graph, with special reference to vocabulary, in directions suggested by the students' insights and by the psychologist.

*Schedule making*, in which they filled in a schedule with class and extra-curricular activities and periods of study in each subject. The students were encouraged to revise the schedules when they did not work out satisfactorily. A supply of mimeographed schedule forms was kept on hand for their convenience.

*Selected Exercises* and S.R.A. Reading Books (20), preceded by practice on short articles from the *Reader's Digest*, as they appeared appropriate.

*The Tachistoscope* demonstrated on two or three occasions according to the time available with periods of practice not exceeding 10 to 15 minutes.

*The Reading Accelerator* demonstrated. Three or four students made use of it when it was assured that it was appropriate. The devices were found to have high motivation value.

*Self-observation* encouraged by the use of an inventory of reading habits, by practice in reading differing textbook material and through lectures and conferences.

*Lectures* and group discussions of the following subjects: Concentration and Distractions; Make Your Own Destiny (a lecture on schedule making); The Principles of Good Reading (3 or 4 lectures); How to Read a Textbook; How to Make an Outline; How to Make a Graph; How to Spell a Word. Few of these lasted in class over 20 minutes or half an hour due to the need for practice periods but were discussed in conference.

The importance of vocational guidance for some students was discussed in relation to concentration problems and the use of the Guidance Center encouraged. The services of the Speech and Hearing Center, and the Eye, Ear, Nose, and Throat Hospital were explained. Eye exercises were available at the Eye, Ear, Nose, and Throat Hospital free of charge provided they were preceded by an ophthalmological examination by a private physician.

## G. RESULTS

### 1. *Some Measurable Results of the Reading Classes*

The Iowa Silent Reading Test was administered the first day and again at the twelfth meeting, using alternate forms and in alternate order. Fifty-one students attended a first and twelfth meeting with an average attendance of 10 meetings. Three students could not remain to complete the seventh part of the terminal test, two took the alternate form some weeks later thus altering the interval between tests, and one student was given the same form twice through a clerical error. Thus the medians are based upon 45 instead of 51 cases of which 35 were men and 9 were women. There was no significant difference between the test results of the men and women and so the measurable results are given for the group as a whole ($t$ value = .721).

The 45 cases gave the following results: (Table 1). The median gain of the total group was 12 standard points from the initial Iowa Silent Reading Test to the terminal test. The median gain was 11 standard points for the 23 freshmen whose scores on the survey test (Coöperative Reading Comprehension Test, *C* 2, Form *Z*) had placed them in the lowest 14 percentiles. The median gain for the other 22 students was 14.5 standard points. These 22 were either freshmen whose survey test results were above the 14th percentile or upperclassmen.

The median standard score of the total group of 45 was 168, corresponding to the 23rd percentile, on the initial test, and 180, corresponding to the 48th percentile, on the terminal test. For the 23 freshmen below the 15th percentile on the Coöperative Test, the initial median standard score was 164 and the terminal score 175, corresponding to the 17th and 37th percentiles, respectively. For the 22 other students, the initial median standard score was 174 corresponding to the 34th percentile and the terminal score 188.5 corresponding to a percentile between the 67th and 70th (Table 2). It would appear from these data that the freshman students above the 14th percentile and the upperclassmen had profited somewhat more than the freshmen below the 15th percentile on the survey test.

Table 2 shows the range of median scores, of percentiles and of gains on the Iowa Test. The means were within one or two points of being identical with the medians and were all significant below the one per cent level (Table 3).

It is recognized that the Coöperative Reading Comprehension Test and the Iowa Test of Silent Reading are quite different tests of reading (3, pp. 497-498; 6, p. 489; 21, pp.480-483; 22, pp. 489-490), and that we have

TABLE 1

RESULTS OF IOWA SILENT READING TESTS ON FIRST AND TWELFTH MEETINGS OF READING IMPROVEMENT CLASSES (MEDIAN SCORES AND GAINS)

| | | Median standard score | | Gain in standard points | Percentile of median standard score | |
|---|---|---|---|---|---|---|
| | No. | Initial test (1st meeting) | Terminal test (12th meeting) | | Initial test (1st meeting) | Terminal test (12th meeting) |
| Freshmen below 15th %ile on Coöperative Reading Comprehension Test | 23 | 164 | 175 | 11 | 17 | 37 |
| Freshmen above 14th %ile on Coöperative Reading Comprehension Test and upper class students | 22 | 174 | 188.5 | 14.5 | 34 | 67-70 |
| Total of all students present for both tests | 45 | 168 | 180 | 12 | 23 | 48 |

TABLE 2

RANGE OF RESULTS OF IOWA SILENT READING TESTS ON FIRST AND TWELFTH MEETINGS OF READING IMPROVEMENT CLASSES

| | | Range of median scores | | Range of gains | Range in percentiles | |
|---|---|---|---|---|---|---|
| | No. | Initial test | Terminal test | | Initial test | Terminal test |
| Freshmen below 15th %ile on Coöperative Reading Comprehension Test | 23 | 152-185 | 164-196 | 2-33 | 6-59 | 14-88 |
| Freshmen above 14th %ile on Coöperative Reading Comprehension Test and upper class students | 22 | 140-196 | 142-208 | 1-26 | 1-88 | 1-99 |
| Total of all students present for both tests | 45 | 140-196 | 142-208 | 2-33 | 1-88 | 1-99 |

measured progress only by the latter. It is likely that to some extent the two tests may be influenced by uncontrollable factors in addition to those usually mentioned. For example, the Coöperative Test is intended to be used principally as a survey test, is administered in large groups and may be expected to be met by the students with varying degrees of motivation. The Iowa Test, in our case at least, is used as further diagnostic aid. Since our classes are non-credit voluntary classes, the motivation on the initial Iowa Test may be expected to be considerably greater than for the mass-administered Coöperative Test. Table 4, therefore, shows a different and more useful distribution of the cases. This distribution it was thought might help to evaluate the group gains somewhat better than that shown in Table 1, since the program was this year designed especially for freshman students most in need of assistance.

TABLE 3
RESULTS OF IOWA SILENT READING TESTS ON FIRST AND TWELFTH MEETINGS OF READING IMPROVEMENT CLASSES

| | Total present for both tests $N = 45$ Iowa Silent Reading Test | | Freshmen below 15th %ile Coöp. $N = 23$ Iowa Silent Reading Test | | Other students taking course $N = 22$ Iowa Silent Reading Test | |
|---|---|---|---|---|---|---|
| | I | II | I | II | I | II |
| Mean | 169.067 | 181.356 | 164.625 | 176.870 | 173.682 | 186.045 |
| *SD* | 12.230 | 12.529 | 8.333 | 8.674 | 13.844 | 14.094 |
| Gain | 12.289 standard points | | 12.218 standard points | | 12.363 standard points | |
| *t* value | 11.999 | | 8.086 | | 8.579 | |
| *P* | .01 | | .01 | | .01 | |
| Median | 168 | 180 | 164 | 175 | 174 | 188.5 |
| Range | 140-196 | 142-208 | 152-185 | 164-196 | 140-196 | 142-208 |
| Gain | 12 standard points | | 11 standard points | | 14.5 standard points | |

The 34 freshmen are divided into four groups, 10 whose scores were below the 15th percentile on both tests and which are dealt with in this paper as the group most in need of assistance, 13 below the 15th percentile on the Coöperative Test but not on the Iowa (there were no cases in which a student fell below the 15th percentile on the Iowa Test but not on the Coöperative Test), eight above the 14th percentile on both, and three who had not taken a Coöperative Test. Upperclassmen were dealt with as a separate group.

A survey of the results of this group of freshmen who were low in both reading tests, of the group who were low on the Coöperative Test and higher

TABLE 4

RESULTS OF IOWA SILENT READING TESTS ON FIRST AND TWELFTH MEETINGS OF READING IMPROVEMENT CLASSES: FRESHMEN GROUPED ACCORDING TO COÖPERATIVE READING COMPREHENSION TEST RESULTS AND UPPERCLASSMEN

| | | Median standard score | | Gain in | Percentile of median standard score | |
|---|---|---|---|---|---|---|
| | No. | Initial test | Terminal test | standard points | Initial test | Terminal test |
| Freshmen below 15th %ile on Coöperative Reading Comprehension and Iowa Silent Reading Test | 10 | 157.5 | 173.5 | 16.0 | 10 | 34 |
| Freshmen below 15th %ile on Coöperative Reading Comprehension Test and above 14th %ile on Iowa Silent Reading Test | 13 | 168.0 | 179.0 | 11.0 | 23 | 46 |
| Freshmen above 14th %ile on both tests | 8 | 184.0 | 195.5 | 11.5 | 57 | 88 |
| Freshmen with no Coöperative Reading Comprehension Test | 3 | 162 | 173 | 11.0 | 14 | 32 |
| Upperclassmen | 11 | 168.0 | 181.0 | 13.0 | 23 | 50 |
| Total | 45 | 168.0 | 180.0 | 12.0 | 23 | 48 |

on the Iowa Silent Reading Test, of the freshmen above the 14th percentile on both tests, those with no Coöperative Test, and of the 11 upperclass students, gives a somewhat different and possibly more accurate picture (Table 4). Since this program was initiated as an effort to give reading assistance on as nearly an individual basis as possible, the small size of these groups does not greatly concern us at this time. Records are being kept on the present groups as well and will continue to be kept as indicators of measurable trends.

The results of the above distribution (Table 4) show a median gain of 16.0 standard points for freshmen below the 15th percentile on both tests, 11.0 standard points for freshmen below the 15th percentile on the Coöperative Test but not on the Iowa Test and for the three without a Cooperative Test, 11.5 standard points for freshmen above the 14th percentile on both tests, and 13.0 standard points for upperclassmen.

The distribution of cases in Table 4 shows the greatest gains (16.0 standard points) among freshmen below the 15th percentile on both tests, a result which might be anticipated since gains are usually more easily made from low than from high scores. If these be real gains instead of artifacts of the test it would appear that the students who derived the most benefit from the classes were the freshmen below the 15th percentile on both tests.

Table 5 presents the range of median scores, of percentiles, and of gains on the Iowa Test of this distribution. For all of the groups, regardless of how we distributed the cases, the gains by *t* test were significant below the one per cent level (Tables 3 and 6).

The norms for the Iowa Test include age and grade norms. The authors of the test state that for college age students the figures lose some of their significance and recommend their use only through the ninth grade. However, we made the comparison, for what it is worth, between the standardization group and those of our group whose initial test scores placed them below the college freshman level. For the freshmen whose scores placed them below the 15th percentile on the Coöperative Reading Test and whose initial scores on the Iowa Test were below the college freshman level the gain was 3.2 grades. For all of the other students enrolled, whose initial Iowa scores were below the college freshman level, the median gain was +3.0 grades (Table 7).

Our lowest group showed an average grade placement of 9.1 on the initial Iowa Test and 12.85 on the terminal test, a gain of 3.75 high school grades in the twelve sessions. The average reading age of this group was 13

### TABLE 5

Range of Results of Iowa Silent Reading Tests on First and Twelfth Meetings of Reading Improvement Classes

| | | Range of median scores | | Range of percentiles | |
|---|---|---|---|---|---|
| | No. | Initial test | Terminal test | Initial test | Terminal test |
| Freshmen below 15th %ile on Coöperative Reading Comprehension Test and Iowa Silent Reading Test | 10 | 152-161 | 164-192 | 6-13 | 17-79 |
| Freshmen below 15th %ile on Coöperative Reading Comprehension Test and above 14th %ile on Iowa Silent Reading Test | 13 | 162-185 | 164-196 | 14-59 | 17-88 |
| Freshmen above 14th %ile on both tests | 8 | 169-196 | 188-208 | 24-88 | 67-99 |
| Freshmen with no Coöperative Reading Comprehension Test | 3 | 160-181 | 169-189 | 12-50 | 24-70 |
| Upperclassmen | 11 | 140-191 | 142-199 | 1-76 | 1-93 |
| Total | 45 | 140-196 | 142-208 | 1-88 | 1-99 |

TABLE 6

RESULTS OF IOWA SILENT READING TESTS ON FIRST AND TWELFTH MEETINGS OF READING IMPROVEMENT CLASSES

| | Total taking first and terminal test | | Freshmen below 15 %ile on Coöp. and Iowa Tests | | Freshmen below 15 %ile on Coöp. and above 14 %ile on Iowa | | Freshmen above 14 %ile on both Coop. and Iowa Tests | | Freshmen with no Coöp. tests | | Upperclassmen | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $N = 45$ | | $N = 10$ | | $N = 13$ | | $N = 8$ | | $N = 3$ | | $N = 11$ | |
| | I Iowa | II | I Iowa | II | I Iowa | II | I Iowa | II | I Iowa | II | I Iowa | II |
| Mean | 169.067 | 181.356 | 157.3 | 173.6 | 170.308 | 179.385 | 183.645 | 196.250 | 167.666 | 177.000 | 168.091 | 181.091 |
| *SD* | 12.230 | 12.529 | 2.948 | 7.200 | 6.533 | 9.026 | 8.645 | 7.395 | 9.463 | 8.641 | 13.839 | 14.701 |
| Gain | 12.289 | | 16.300 | | 9.077 | | 12.625 | | 9.333 | | 13.000 | |
| *t* value | 11.999 | | 7.193 | | 5.984 | | 5.905 | | 10.594 | | 3.185 | |
| *P* | .01 | | .01 | | .01 | | .01 | | .01 | | .01 | |
| Median | 168 | 180 | 157.5 | 173.5 | 168.0 | 179.0 | 184.0 | 195.5 | 162 | 173 | 168 | 181 |
| Range | 140-196 | 142-208 | 152-161 | 164-192 | 162-185 | 164-196 | 169-196 | 188-208 | 160-181 | 169-189 | 140-157 | 142-199 |
| Gain | 12.0 | | 16.0 | | 11.0 | | 11.5 | | 10 | | 13 | |

TABLE 7

RESULTS OF IOWA SILENT READING TESTS ON FIRST AND TWELFTH MEETINGS SHOWING MEDIAN GAINS IN GRADE PLACEMENT*

| | Total | | Freshmen below 15th %ile Coöp. Test | | Others | |
|---|---|---|---|---|---|---|
| | $N = 30$ | | $N = 19$ | | $N = 11$ | |
| | Iowa Silent Reading Test | | Iowa Silent Reading Test | | Iowa Silent Reading Test | |
| | I | II | I | II | I | II |
| Median standard score | 161.5 | 174.0 | 161.0 | 174.0 | 162.0 | 178.0 |
| Median grade placement | 9.9 | 13.0 | 9.8 | 13.0 | 10.0 | +13.0 |
| Gain in school grades | 3.1 | | 3.2 | | +3.0 | |

*Includes only those students whose initial Iowa Test scores placed them below the 13th grade.

years 1 month on the initial test, and plus 18 years 2 months on the terminal test (Table 8). The norms do not extend higher than 18 years 2 months.

The average gain for all students whose initial Iowa Test scores placed them below the 13th grade or college freshman level was 3.1 school grades (Tables 7 and 8).

Some improvement from practice effects must be considered in all of these results even though a different form of the test was given at the end of the course. A true control group would have been desirable. During the second year of the reading program we tested older freshman students in two English classes one month apart on the alternate forms, *Am* and *Bm*, of the Iowa Silent Reading Test. This was done as a matter of interest rather than because any significant value could be attached to controlling only the two factors, freshman status and attendance in English classes. Thirty-one students took the initial test and 18 or 58 per cent were present for the terminal test.

TABLE 8

RESULTS OF IOWA SILENT READING TESTS ON FIRST AND TWELFTH MEETINGS SHOWING MEDIAN GAINS IN GRADE PLACEMENT FOR STUDENTS BELOW 13TH GRADE ON INITIAL IOWA TEST

| | Total $N = 30$* Iowa Silent Reading Test I | Total $N = 30$* Iowa Silent Reading Test II | Freshmen below 15th %ile Coöp. and Iowa Tests $N = 10$ Iowa Silent Reading Test I | Freshmen below 15th %ile Coöp. and Iowa Tests $N = 10$ Iowa Silent Reading Test II |
|---|---|---|---|---|
| Median standard score | 161.5 | 174.0 | 157.5 | 173.5 |
| Median grade placement | 9.9 | 13.0 | 9.1 | 12.9 |
| Gain in school grades | 3.1 | | 3.8 | |

*Includes only those students whose initial Iowa Test scores placed them below the 13th grade or college freshman level.

The median gain for the 18 cases was seven standard points as compared with a gain of 14 standard points for the 34 freshmen who had taken the reading course. The median for the English classes was 174 or the 34th percentile on the initial test and 181 or the 50th percentile on the terminal test (Table 9). For the reading classes the freshman median was a standard score of 166 or the 20th percentile on the initial test and 179 or the 46th percentile on the terminal test.

The range in standard scores of the English classes was from 152 to 200 on the initial test and from 150 to 206 on the terminal test. In percentile the range was from the 6th to the 95th on the initial test and from the 4th to the 99th on the terminal test.

For the 34 freshmen who attended reading classes the range in standard scores was from 152 to 196 or from the 6th to the 88th percentile on the initial test and from 162 to 208 or from the 14th to the 99th percentile on the terminal test.

The "*t*" values for the control group were also significant below the one per cent level as in the case of the reading classes (Table 9). This may be a function of the facts that the control students were older students and with one exception did not consider that they had reading problems.

TABLE 9

RESULTS OF IOWA SILENT READING TESTS ON FIRST AND TWELFTH MEETINGS OF READING IMPROVEMENT PROGRAM: MEDIAN STANDARD SCORES ON IOWA SILENT READING TESTS OF CONTROL GROUP

$N = 18$

| | Total present for both tests $N = 18$ Iowa Silent Reading Test I | II | Class I N = 12 Iowa Silent Reading Test I | II | Class II $N = 6$ Iowa Silent Reading Test I | II |
|---|---|---|---|---|---|---|
| Mean | 171.389 | 179.222 | 173.917 | 182.500 | 166.333 | 172.667 |
| *SD* | 13.075 | 16.207 | 13.183 | 16.059 | 11.270 | 14.403 |
| Gain | 7.833 standard points | | 8.583 standard points | | 6.334 standard points | |
| *t* value | 4.466 | | 3.801 | | 2.906 | |
| *P* | .01 | | .01 | | .01 | |
| Median | 174 | 181 | 174.5 | 182.5 | 163.5 | 172.0 |
| Range | 152-200 | 150-206 | 154-200 | 150-206 | 152-184 | 153-192 |
| Gain | 7 standard points | | 8 standard points | | 8.5 standard points | |

The number of assignments and exercises completed did not appear to be highly related to over-all gains on the Iowa Test. The median number of completed assignments and exercises for the entire group was 16. The gain in standard points for this group was 12. The median number of assignments and exercises completed by the freshmen below the 15th percentile both on the Coöperative and Iowa tests was 13 whereas the gain in standard points was 16. The correlation (Pearson product-moment) between the number of exercises and the gains made by the entire group was —0-11. This is not significantly different from zero, and the inference may be drawn that there is no relationship between the number of exercises completed and the gains made by the group.

College aptitude as measured by the Ohio Scholastic Aptitude Test showed no relationship to the over-all median gains on the Iowa Silent Reading Test on 23 students for whom we had data on both tests. The Ohio Test is administered at the beginning of the freshman year and with one exception these 23 students were freshmen. A rank-order correlation between the Ohio Test and the gains on the Iowa Test was .005.

Other rank-order correlations computed on this group of 23 students were as follows:

Ohio Scholastic Aptitude Test and Coöperative Reading Comprehension Test, *C* 2, Form *Z*, $r = .82$.

Ohio Scholastic Aptitude and initial Iowa Silent Reading Test, $r = .75$.

Coöperative Reading Comprehension Test, *C* 2, Form *Z*, and initial Iowa Silent Reading Test, $r = .75$.

It is of interest that four freshman students whose scores on the Coöperative Test placed them in the lowest 15 percentiles were on deans' lists of honor students before the end of the year. While this cannot be assigned to the reading classes it is doubtful if they could have attained this without improvement of their reading, provided the tests were an accurate measure of this. Of these four, one student (No. 38) was below the 15th percentile on both the Coöperative and Iowa tests. While the other three were below the 15th percentile on the Coöperative test their initial Iowa scores placed them at the 17th, 24th, and 37th percentiles respectively. Again this raises the question as to the extent to which the former test represented their reading. One of these students was under tension to the extent of needing medical advice, a condition for which one must be alert in reading problems. In our experience a true reading problem does not occur as an isolated problem but as a symptom of other more basic problems. For this reason caution is necessary in dealing only with the symptoms, as for example through "speeded" reading and the unrestricted use of accelerating devices.

Two students were denied re-admission the following year by reason of academic failure, both of whom were below the 15th percentile on the Coöperative Test. Only one however, was below that level both on the Coöperative Test and the initial Iowa Test.

The vision of 19 students was tested by means of a Keystone Visual Survey Telebinocular. It was noted that some muscular imbalances were common, a finding which has been reported more completely by other investigators (12, 17). Clinical experiences with visual factors in relation to personality have led us to believe that visual factors are not causes of reading difficulties but rather that they are results of a more basic cause or causes which will be discussed later in this study (8, pp. 141-148; 12, p. 120; 15, pp. 7-33, 100; 18, p. 17).

We are aware of the inadequacy of testing for complete and accurate measurement of progress. It is clearly stated to the students that testing is only one aid both in the diagnosis of reading difficulties and in the estimate of progress. However, the Iowa Silent Reading Test is one which is in common use and has the advantage of providing some measurable bases of comparison (1, 23). We are further aware of the disadvantages of a single measure over a more complete testing program for the measurement of progress. This was not expedient in a new program which was designed this year principally to assist those students who, because of reading disadvantages, had the least amount of time to give to a non-credit course. Also, we share to some extent the opinion of the late Walter B. Cannon:

> It is a satisfaction to know that the eminent physical chemist, B. N. Lewis, has declared, "I have no patience with attempts to identify science with measurement, which is but one of its tools, or with any definition of the scientists that would exclude a Darwin, a Pasteur, or a Kekule." To these three may be added Harvey, Virchow, Pavlov, Sherrington, and many others (5, p. 36).

In interpreting the results of the testing program, however, as a measure of progress it is to be kept in mind that the instruction periods of these classes, excluding the two periods devoted to administration of alternate forms of the Iowa Silent Reading Test, were 10 50-minute periods, supplemented by conferences to the extent which time permitted both with respect to the number of students and the time which they had available during and outside of class periods. Can the reading gains made in one month be accounted for solely by instruction? (Tables 1-10.) For example, can we assume that a student below the 15th percentile on both the survey and diagnostic tests progress in 12 meetings from the 11th to the 79th percentile on the Iowa Test (Table 10) unless factors other than reading instruction were operative? (12, p. 114). We believe that these factors may be expected to be physical and psychological factors in combination and we are aware that only for purposes of discussion can the two be separated. This discussion will confine itself to some psychological factors which seemed to precipitate out from our observations of the young people who comprised the group.

### 2. *Non-Measurable Results of the Reading Program*

Clinical observations do not lend themselves to quantification at present. Yet, in accounting for the factors other than instruction which are operative in bringing about measurable results of reading improvement which appear to be somewhat out of proportion to the amount of instruction possible in one month or 12 class sessions, one must rely to some extent upon observation and clinical insights. Observations noted on these 45 students, reinforced by those made over a period of years upon children and young adults with reading problems have led to some opinions and some hypotheses with respect to problems of reading.

TABLE 10
RESULTS OF IOWA SILENT READING TEST ON FIRST AND TWELFTH MEETINGS OF READING IMPROVEMENT CLASSES

$N = 45$

| | Median standard scores | | Percentile rank | |
|---|---|---|---|---|
| | Initial | Terminal | Initial | Terminal |
| 1. | 196 | 197 | 88 | 90 |
| 2. | 193 | 204 | 82 | 98 |
| 3. | 191 | 196 | 76 | 88 |
| 4. | 189 | 203 | 70 | 98 |
| 5. | 187 | 208 | 64 | 99 |
| 6. | 186 | 193 | 62 | 82 |
| 7. | 185 | 196 | 59 | 88 |
| 8. | 181 | 192 | 50 | 79 |
| 9. | 181 | 190 | 50 | 70 |
| 10. | 181 | 189 | 50 | 73 |
| 11. | 179 | 190 | 46 | 73 |
| 12. | 179 | 184 | 46 | 57 |
| 13. | 175 | 188 | 37 | 67 |
| 14. | 175 | 180 | 37 | 50 |
| 15. | 174 | 186 | 34 | 62 |
| 16. | 173 | 199 | 32 | 93 |
| 17. | 173 | 189 | 32 | 70 |
| 18. | 171 | 184 | 28 | 57 |
| 19. | 169 | 188 | 24 | 67 |
| 20. | 169 | 168 | 24 | 23 |
| 21. | 168 | 187 | 23 | 32 |
| 22. | 168 | 177 | 23 | 42 |
| 23. | 168 | 173 | 23 | 64 |
| 24. | 167 | 178 | 22 | 45 |
| 25. | 165 | 174 | 19 | 34 |
| 26. | 165 | 169 | 19 | 24 |
| 27. | 164 | 179 | 17 | 42 |
| 28. | 164 | 177 | 17 | 46 |
| 29. | 162 | 173 | 14 | 46 |
| 30. | 162 | 164 | 14 | 17 |
| 31. | 161 | 174 | 13 | 34 |
| 32. | 160 | 176 | 12 | 24 |
| 33. | 160 | 174 | 12 | 39 |
| 34. | 160 | 169 | 12 | 34 |
| 35. | 160 | 166 | 12 | 20 |
| 36. | 159 | 181 | 11 | 50 |
| 37. | 159 | 178 | 11 | 45 |
| 38. | 158 | 192 | 10 | 79 |
| 39. | 157 | 180 | 9 | 50 |
| 40. | 157 | 173 | 9 | 32 |
| 41. | 156 | 175 | 8 | 32 |
| 42. | 156 | 173 | 8 | 37 |
| 43. | 153 | 169 | 6 | 17 |
| 44. | 152 | 164 | 6 | 24 |
| 45 | 140 | 142 | 1 | 1 |

The first opinion has to do with terminology. While it is generally agreed among instructors of reading that one of the first principles of good reading is clear-cut definition of terminology with which to formulate accurate concepts, the literature upon the subject of reading problems is replete with terms which are vague and can, therefore, lead only to vague concepts with respect to the problems and their solution. We shall attempt, therefore, to avoid such a general term as "emotion" and replace it with the name of a specific emotion when appropriate and possible.

The mere matter of that courage which comes with living under college conditions and succeeding in its demands is believed to influence reading test results regardless of training. Various terms are employed to describe this process, as for example maturation, independence, and ego development (2, 4, 9, 16), but it is our belief that regardless of terminology a strengthening of the individual takes place which amounts to the attainment of courage. In addition to this positive acquisition, a fear or several fears have possibly been replaced by the opposite emotion, courage, through successful experiences. This, we believe, is largely responsible for the improvement of the control groups in the Dartmouth study (12, p. 114, 6a) as well as of many students in college everywhere. Many students are afraid of not succeeding, afraid of being out on their own, away from their parents, although they might deny this because they do not always recognize it nor can they afford to, they often believe. It is actually not necessary to bring these fears to a conscious level in many cases. Successful experience within their own definitions of success is sufficient to lead to the necessary courage. No doubt there will be those who object at first to the term courage as opposed to emotional maturation or to independence or ego development but we believe that analysis of the latter terms will finally lead to the former.

Discretion is necessary in discussing recent cases and for this reason case studies cannot be reported in detail in this study. More appropriately at this time are the observations summarized and discussed from the point of view of their possible dynamics in the reading situation.

## H. OBSERVATIONS

Whenever a personality factor is discussed it would be desirable to know and to state the degree to which this factor is present as compared to a standard. In the case of students with no reading handicap we may assume that if the personality factors are similar in kind to those of students with a reading handicap they are at least in a state of equilibrium in the former whereas the same personality factors are not in a state of equilibrium but may be assumed to be in excess in positive or negative amounts in those students with reading problems. So far we can only safely describe personality factors in terms of behavior, or at a descriptive level and decide somewhat subjectively whether the behavior is similar to, in excess of, or less in amount than in an equally subjectively determined standard. To call the standard group "a control group" is common practice but is in many instances a questionable one in our opinion. Human variables are too many and the relationships too complex, and much as we might wish to simplify them, saying that we can and proceeding as though we are able to, only leads us farther away from the facts instead of closer to them. It is even dangerous because we behave as though we had measurable data to go on.

The following observations are in no way unusual and are met and recognized by those who work with reading problems. In discussing them with the students, however, we were led to a somewhat different interpretation from those generally made relative to College-age students.

The students' approach to reading appeared to characterize to some extent their approach to life situations generally. Those who found it difficult to attack a single word, which in this reading program represented only two students and to whom it was necessary to give individual instruction because of extreme disability, appeared to have difficulty also in getting started or in beginning things generally. It is of interest that these two students compensated for their reading handicap in two different ways. One was highly conventional and rigid in his behavior and would not attempt a word unless he felt reasonably certain of it. The other assumed an attitude of extreme bravado and nonchalance and called a word anything which came to mind with little regard for its real form. Both of these kinds of behavior are familiar to the reading clinician. Both students were afraid to begin and met their problem according to characteristic patterns of behavior.

The difficulty in approaching sentences, paragraphs, longer passages and even books, likewise appeared to characterize the student's approach to life situations with chiefly a matter of degree differentiating him from those whose word attack was inferior. "I'm afraid to start things" was not an unusual comment. Those who reported a good start with diminishing interest after starting, not infrequently reported, "I never finish anything" or "I have to make myself finish things."

In those instances in which the students dropped out, it was not unusual for them to make such comments as: "I just don't have time to put in. When I do something I like to put a lot into it and I don't have the time to put anything much into this extra course." These students while not highly motivated in improving reading were highly motivated in other directions and doing well in subjects in which reading was minimal as, for example, in music and art. They did not want to participate in an activity in which they could participate but superficially.

On the other hand, the students who dropped out when they found that work and effort were involved and that mechanical devices alone would not improve their reading, were found to be somewhat superficial in their approach to situations in general. They were students who were far more interested in social contacts than in academic work. Three of them wanted only to work with devices and to forego even the diagnostic tests. These students were "rugged individualists," exhibited a high degree of rigidity and wanted "to set the rules for the game or not play." The fourth was thoroughly coöperative until it was necessary to put forth effort in academic work. All of them indicated an unusual need for success in social situations and preferred social over academic activities.

Over and over it became apparent that each student's reading is something of a cross-section of his behavior in life situations. Because of the cultural importance of reading and because success and failure in this society hinges to a considerable extent upon success in reading, conduct in the reading situation is strongly high-lighted. Hence also comes part of a fear attached to reading since failure in reading is almost tantamount to failure in life in our society.

Obviously one must ask several questions at this point: Why do some students have more fear of the reading situation than others? Why do some students have reading difficulties and why do some not have reading difficulties even though they may have the same personality traits described above which accompany reading problems (e.g., difficulty in starting, giving up in varying degrees, etc.)? These and other questions must be answered if we are to understand the nature of individual reading problems and in every instance the reading group is made up of individuals who must be dealt with as such. In our observations of these young people of both sexes whom we knew and treated as individuals, we perceived some characteristics which may shed light upon these questions.

The students with reading problems in general fell into two extreme categories of behavior: docility and mischievousness. For the most part they were too compliant, ill-confident, and wanted much more direction than their counterparts in the college population. They were afraid to venture and tended to cling. Even though some of them might indulge in mischief such as hang out of windows, call out to other students before class, and make "smarty" remarks they revealed their dependence upon external controls which were, for the most part, not forthcoming in this clinical setting. Few of them were daring enough even in this permissive atmosphere to be outright defiant. They could come or go as they pleased or not attend classes since they were told that absences were not reported. The tendency was to stay after class either to complete an exercise or to do an extra one, come early to class, come for extra work, or merely to come to the rooms for study. Several of them remarked about the "peaceful, quiet atmosphere" although the rooms were over the dance studio. When they did not attend class, they would usually either apologize or telephone if they could not come and almost always would volunteer a reason in response to which the psychologist would say that they had been missed and show pleasure in their return. In most instances they would ask to make up the work.

Since few demands were made they soon discovered that what they accomplished depended to some extent upon their initiative and motivation as they had been told in the beginning. They were encouraged individually as much as time permitted with praise, suggestions, and warm interest. The psychologist never allowed them to become dependent nor did she reject any of them. Some of the most timid ones gradually came to venture to laugh at the "show off" behavior of others and in a few instances joined in. In one instance at least this was a real victory for the student. He had been somewhat depressed upon entering the class, came sullenly in each day, sat by himself and worked. Only towards the end of the course did he gather his courage to approach the psychologist unsolicited. One exceptionally timid boy was allowed to come to a second series of classes because he was beginning to "expand" and gathered courage at the end of the first series to ask if he might return. It was not always possible nor desirable to do this when students requested it but in this instance he was allowed to return. By the end of the second series he was a somewhat changed person in that he had become more out-going, showed a sense of humor and greatly increased courage.

Success and confidence obviously worked together. As the students found themselves independent individuals who were respected and encouraged but not coerced, as they learned and put into practice the suggestions made during the lectures, they quite apparently relaxed in their attitudes towards reading. It was satisfying to hear their comments: "I'm getting a lot more out of my chemistry now. It really makes sense and it doesn't take me nearly as long as it used to." "It's the funniest thing, I'm getting better in history. I used to hate it." "I'm really getting some place now."

On the other hand, there were those whose progress was negligible, if any. We were constantly reminded of the disadvantages of large group work for some students, of the need for more time for those students whose experiences with reading were especially limited and whose confidence in themselves with respect to reading was so low that we were unable to reach to its depths in this program. With smaller groups and more individual work we might have succeeded more often in helping the students who particularly needed individual encouragement. This was particularly true of Students 21, 29, and 45. Nos. 21 and 45 needed a great deal of individual work but were unable to do this and eventually were refused re-admission to college. No. 29 maintained honor roll status at heavy cost by reason of excessively neurotic drive.

In summary, the evidence was cumulatively in the direction of an increased confidence or courage towards approaching that which had previously offered a threat. There is nothing new in this observation, however, as it is a well-known fact that one does better that which he feels he can do than under the opposite condition. The question now becomes more specific as to why the majority of these students were handicapped or felt handicapped in reading.

Superficially, one might say that a lack of reading experiences might be a cause, but we must look beyond to the reasons for lack of reading experiences. Family interests and cultural background are not sufficient causal factors since many a highly educated person has come from a family in which little reading was done and from homes in which there were few books to be read. All of these students had been "exposed" to books and reading through a high school education and were admitted to the university on a selective basis. While family interests may not have promoted an interest in reading in some instances, it was too often remarked: "My mother (or father) are great readers and I cannot understand why I have such a hard time reading," or "My folks don't know what to make of me. They read a lot, especially my father (or mother)," and other similar statements.

One common factor which appeared even in the better students was the ever present but usually (by the student) unrecognized factor of fear. As uncomfortable as this truth may be, once it is accepted by parents, teachers, and therapists, the process of teaching reading is simplified. This fear has its roots in many directions.

Physically it may be partially rooted in the too early requirement of reading. Optically some children are required to make the fine coördinations necessary to reading before the eye muscles are sufficiently mature to perform this act (11, pp. 129-136). Binocular coördination is an essential to good reading (15, p. 62). If this is not possible because of physical immaturity and the children are required to make the effort to read they can only meet with frustration which in turn is known to have a disturbing effect upon vision (7, pp. 352-368).

When failure is encountered even because of a simple lack of physical development it is still a failure to the child and the parents. Regardless of the rationale, if a child cannot perform an act which has been established as necessary he is regarded as a "problem" by his teacher, parents, and something of a failure by himself. Thus we help to create reading problems where they could perhaps have been avoided if the stage of development of the eye muscles, which make coördination possible, had been determined before the child began to read.

If, however, there is no question of physical development, there may still be a reading problem on a psychological basis. Reading is a psycho-physical act. Just as the physical act of sight may in itself be interfered with under certain psychological conditions, as for example by fear (7, pp. 352-368) so the complete reading process may be similarly inhibited. Let us trace this factor to its logical sources and we shall arrive at the same sources at which we seem to arrive invariably in clinical practice.

Reading is an act which implies a certain amount of "growing up." When a child who enters school is too fearful of the process of growing up, by reason of his incapacity to meet the social and emotional requirements inherent in the term, he has become this way by reason of his experiences up to this time. Those children who cannot depend upon themselves but who must depend upon others to an extent beyond a reasonable expectance for their years, have for the most part depended upon their mothers. This does not preclude the possibility of their having depended also upon their fathers. One parent more dominant than the other is selected by the boys whereas the less dominant parent is more often selected by girls for the greatest degree of identification in the traditional North American culture. Thus, if a boy is dependent upon his mother unduly, she is likely to be the dominant parent and the one with whom he most identifies himself, not, of course, with conscious intent. Since his mother cannot go to school with him, the teacher becomes the substitute for the mother. Dependence begets fear, and the boy may transfer his fear to his teacher of reading and to the reading situation under circumstances which are discussed later. Meanwhile the teacher attempts to teach him to read with negative results in extreme cases. If she would devote her full time to him he might possibly learn to read, but this is not realistically possible nor even desirable. His need is to be helped towards independence in a kindly understanding way. Fostering dependence fosters a potential reading problem if not a present one. In the same way a too-authoritative teacher can create a reading problem where one did not previously exist since all children are potentially dependent and the degree to which this condition is fostered is a critical factor in both development and reading.

This dependence-identification-fear syndrome in reading problems is demonstrated through a comparison of the performance in reading with that of mathematics. Some diagnosticians use mathematics as one base for determining whether or not a reading problem exists as contrasted with a problem in general intelligence or academic aptitude (10, p. 214). Students having true reading problems typically perform better in mathematics which is a subject more often identified in our culture with males than with females. Therefore, does it not appear reasonable that the child readily learns arithmetic from the same teacher who unsuccessfully tries to teach him reading since he, the child, is not overly dependent upon his father, cannot identify himself *strongly* with his father, is not afraid of his father, and consequently has no barriers to learning that which is associated with his father? If the teacher be a man, the reverse might be true. The boys would more often have difficulty in mathematics, and read better, provided that the family identification were stronger with the father as the more dominant parent, and the teacher were identified with this parent. The hypothetical completely balanced or homeostatic masculine person is, therefore, on these premises not to be expected to have reading as his greatest interest. He may be expected also to have no reading problem. Our findings in clinical practice have borne out these statements.

The girls, on the other hand, who normally identify more with the less dominant parent, the mother, are found to identify with the father when he is the less dominant parent. This leads to rather better performance in mathematics than in reading. Women mathematicians will be found to come within this group of father identified women. The importance of the identification process in children to all forms of learning (13, 14) and perceptual processes may readily be seen.

We believe that it may be safely stated at this time that the degree to which a child identifies himself with each parent determines to a considerable extent his success or its opposite in activities associated with each parent. More specifically our findings suggest very strongly that reading problems occur with those students whose identification is more strongly or perhaps excessive with the parent who is the more dominant and provided that this parent is the one more interested in reading. In the school situation the sex and personality of the teacher either reinforce or act as therapeutic agents in this established family constellation as its affects the child's learning (14, pp. 219-225). Obviously, it must always be a matter of degree with respect to all of the elements involved including identification and thereby dominance and fear.

We have intentionally omitted a discussion of the important factor of relatively unconscious anger towards the dominant parent in the syndrome of reading problems. The frustrations to a child or young person of a dominating, authoritative parent, however charmingly disguised, and the resulting anger towards the parent are relatively unrecognized by those who have reading problems. Some popular recognition of the effects of anger upon *vision* are apparent however through such folk-sayings as ". . . so angry I can't see straight." We believe that a state of relatively unrecognized anger as well as fear will be found to account in part for the muscular imbalances often associated with reading problems (7, 8, 11, 12, 15, 17, 18).

It is not, in our opinion, however, as necessary to deal with the anger *per se* as with the fear factor. Once the fear or fears are relieved, the young person is more capable of dealing with frustrating situations himself and thus he diminishes the cause of anger. We do not, however, wish to appear to underestimate the importance of the anger factor as part of the protest against identification and hence against reading, but rather to call attention to the conditions surrounding the fear factor as perhaps primary. In practice we deal with both through the permissive situation, open discussion, and to a somewhat limited extent through conference.

## I. SUMMARY

The present report has attempted to describe briefly a college program for the improvement of reading, the results of which are under continuing investigation. It was conducted by the writer, with the assistance of a secretary. The freshman survey test was administered by Dr. Cecil W. Mann, of the Psychology Department of Tulane University, with the assistance of graduate students, and was scored under the direction of Mr. Robert Long and the writer.

While a number of different services were provided the first semester which assisted in eliciting the interest of the students in a reading improvement program, the classes herein described began the second semester of the college year 1953-54.

The classes consisted of 12 50-minute periods supplemented by informal conferences within the time limits imposed by student and clinician schedules. One hundred and three students enrolled. The course was voluntary, carried no credits, and no penalty for non-attendance in class or withdrawal. Fifty-one students attended a first and twelfth meeting with an average class attendance of 10 meetings. Seven of these cases had to be excluded from the statistical study due to the following: Three girls had to leave to meet another class before the end of the second test and did not return to complete it. Three boys did not take the test on the twelfth meeting but returned several weeks later thus altering the interval between tests; one student was given the same form twice through a clerical error. Thus the experimental group consisted of 45 students, 36 of whom were boys and 9 were girls.

Of the 45 students, 23 were freshmen whose test results on the Coöperative Reading Comprehension Test given to the freshman class in the fall were below the 15th percentile. Eight of these students were girls. Twenty-two other students completed an initial and terminal test as well, only one of which was a girl. There was no significant difference between the test results of the boys and the girls and so the measurable results are given for the group as a whole ($t$ value = .721). Ten students were below the 15th percentile on both the Coöperative Reading Comprehension Test and the Iowa Silent Reading Test, and are dealt with as the group most in need of assistance.

These students are admitted to Tulane University by criteria which are determined by the University Administration on the basis of scholarship and character and would, therefore, be representative of men's and women's colleges in which a similar policy obtains.

The methods employed were the giving of actual assistance in the reading process itself together with suggestions for practical ways in which to adjust to the college situation. Group discussion and limited conferences related to the perplexities of students were included in the course. Psychotherapeutic techniques were employed through a permissive situation in which the responsibility chiefly rested with the students for work accomplished and in which the encouragement and confidence of the clinician played a major rôle. Observations and clinical insights are the basis of some of the conclusions of the study of these students who sought help in reading.

The Iowa Silent Reading Test was administered at the first meeting and again at the twelfth meeting in alternate form. Approximately one-half of the students were first given Form *Am* and later Form *Bm* and the other half in opposite order. A summary of the measurable results is given below and is followed by a summary of the non-measurable results and finally, by some conclusions and hypotheses.

## J. SUMMARY OF RESULTS

### 1. *Summary of Measurable Results*

1. Over-all median gains from the first to the twelfth meetings varied from 11.0 to 16 standard points, on the Iowa Silent Reading Test, according to statistical groupings. The means were within one or two points of being identical with the medians and by the "*t*" test for matched groups all were significant below the one per cent level.

2. The over-all median gain for all students was 12 standard points.

3. The greatest median gains, 16 standard points, were among the students showing greatest need, i.e., among students whose initial test scores placed them below the 15th percentile on both the Coöperative Reading Comprehension Test and the Iowa Silent Reading Test.

4. The next largest median gains, 13 standard points, were among the students who were above freshman standing.

5. The least median gains, 11.0 standard points, were among freshmen who were above the 14th percentile on the Coöperative and Iowa Tests, and below the 15th percentile on the Coöperative and above the 14th percentile on the Iowa, respectively, and three freshmen having no Coöperative Test.

6. There appeared to be no relationship between the Ohio Scholastic Aptitude Test and the over-all median gains of 23 students for whom there was data on both tests. The rank-order correlation was .005.

7. Rank order correlations between tests for the same group were: .82 between the Ohio Scholastic Aptitude Test and the Coöperative Reading Comprehension Test; .75 between the Ohio Scholastic Aptitude Test and the initial Iowa Silent Reading Test; .75 between the Coöperative Reading Comprehension Test and the initial Iowa Silent Reading Test.

8. The largest gains were not necessarily among those students who completed the most assignments.

9. While relatively few students [18] were tested on the Keystone Visual Survey Telebinocular it was noted that, by this test, muscular imbalances were somewhat common, a finding frequently reported by other investigators.

Four freshman students whose scores on the Coöperative Reading Comprehension Test were in the lowest 15 percentiles were subsequently on deans' lists of honor students. This cannot be assigned to the reading classes although it is doubtful if they could have attained this without improvement of their reading, provided of course that this test was a measure of their reading. In at least three cases this is doubtful. One of these students was below the 15th percentile on both the initial Iowa and the Coöperative Tests. None was above the 30th percentile on the initial Iowa Test. Seven of nine upperclassmen who were not graduates were also on deans' lists of honor students. Since all of these students elected to take the course it would be reasonable to assume that the motivation which led them to do so corresponded to that which might lead them also to unusual academic achievement.

Two students were refused re-admission to the University the following year by reason of academic failure. One of these was below the 15th percentile on both the Coöperative Test and initial Iowa Test. Both had problems beyond the scope of this program.

### 2. *Summary of Non-Measurable Results*

1. A marked increase in self-confidence or courage was noted in the students who also made gains in standard points. It is not possible to measure these advantages at present but they appeared to be somewhat in proportion to measurable gains.
2. Reading problems provided to some extent a cross-section of the individual's general behavior.
3. Students gained most in reading who appeared to gain most in self-observation provided they were able to translate their observations into appropriate action. This implies that in addition to intellectual insight there had to be emotional change.

Up to this point we have introduced, perhaps, little that is new to the clinician who is familiar with reading problems. Psycho-analytic and psychological literature contain studies and reports of clinical cases which are in agreement with our findings up to a point (9, pp. 285-289; 4, pp. 132-137; 16, pp. 114-153; 14, p. 100). Identification (13, pp. 322-386; 14, pp. 219-225), immature ego development (2), vision (16, p. 98), and fear (18) have all been suggested as factors in reading problems. Stating these observations, however, is somewhat the equivalent of a doctor's telling a patient the name of his illness unless one gets even closer to the basis of the difficulties. This latter is necessary in order to apply appropriate remedial measures. To do this it will be necessary to weave our findings and those of other investigators into an appropriate whole through the orderly development of an appropriate theory of learning based upon an appropriate concept of emotion, both as free from "higher order abstractions" as psychologists can make them. Even this has been partially done for us throughout history so that what we contribute is perhaps a new synthesis or perhaps only a different orientation. This follows in the conclusions and hypotheses based upon clinical experiences and more recently upon work with college age students who either were inferior in reading according to standards for the tests employed, or who felt inferior, or who merely wanted to improve their present reading status. For 45 of these we have included some test data.

The following conclusions and hypotheses are presented with respect to reading as an example of some perceptional and learning processes, themselves major forms of human behavior and conduct.

## K. Conclusions and Hypotheses

1. This study demonstrates that as far as the reading process is concerned, as measured by the Iowa Silent Reading Test, short intensive practical courses within our population, conducted with psycho-therapeutic techniques, lead to significant measured changes which compare favorably with traditional courses lasting a semester.

2. Gains at the college level of a comparable population are believed to be possible in part as a function of group psycho-therapy supplemented by short conferences directed in part towards individual problems of which reading problems are by-products.

3. The measured duration, direction, and transfer effects of the gains are not known at this time.

4. Median gains among the experimental group were 12 standard points compared with 7 standard points among older groups of freshmen in two English classes selected by the English Department. The mean gains on this group were also significant below the one per cent level. While this was not a proper control group excepting with respect to taking English and being freshmen it made an interesting comparison in favor of some of the hypotheses which follow.

5. On the basis of the results of the older group of freshmen we may conclude that by the Iowa Test older students with reading problems who make seven points of gain show significant improvement.

6. Reading as a perceptual process first, and as an adaptive or learning process next, is interfered with by a fear factor.

7. Reading problems are symptoms and as such respond to some extent to symptomatic treatment but the real causes lie embedded in the early life experiences of the individual which have created a fear factor which becomes associated with the act of learning to read.

*a*. The fear factor is originally related to the parent with whom the child identifies himself more strongly, and possibly excessively.

*b*. If this identification be with the parent more interested in reading, a reading problem is incipient and may develop provided that:

(1). The personality of the teacher is sufficiently similar to the parent with whom the child most identifies himself.

(2). This parent creates dependence in the child to an exaggerated degree.

(3). Resistance to reading represents in part an effort to attain or to maintain individual integrity.

*c*. Because of dependency inherent in childhood the fear factor may possibly be initially introduced by the teacher or by the school.

*d*. We believe that some generalization of the effects which lead to fear and consequent reading problems was demonstrated, and that "effects of generalization" may be substituted as an alternate term for identification in this study.

*e*. Treating symptoms alone must be questioned. The symptoms may be in the interests of the entire bodily economy. This includes somatic as well as psychological economy of the individual.

8. The degree to which the fear factor is able to be reduced in the learning situation determines in part the degree of reading improvement or learning.

9. The degree of disability is somewhat in proportion to the number and intensity of fear-producing experiences. This carries with it the implication that the degree of disability is for example, somewhat in proportion to the amount of dependence which leads to fear.

10. The nature of the disability or reading problem is dependent upon the nature of the fear-producing experiences. For example, some children read music and/or mathematical symbols who have difficulty in reading the symbols which make words and the word symbols which make sentences, etc. Also, some students have problems of initial attack while others begin but are inhibited in varying distances along the way, behavior which may be related to early feeding and possibly to early toilet-training experiences.

11. The valence of fear-producing stimuli is specific to each individual and whether there is a reading problem or other problem without an accompanying reading problem is determined by specific factors the nature of which we believe that we have in part identified in Nos. 7 and 10, and which warrants continued investigation.

*a*. Reduction of the fear factor allows the subject freer action against frustrations with a resulting reduction of anger concomitants.

(1). We believe that a hierarchy of emotions, as they are related to reading problems, was demonstrated among our population with fear taking precedence in order of need for treatment.

(2). Anger concomitants are not to be disregarded in therapeutic reading situations at the college level but are considered at this time to come second to fear reactions in treatment, although they may be primary in origin.

12. If these hypotheses be true it is logical to assume that visual problems in general, and more particularly for our present purposes those associated with reading problems (e.g., aneisikonia and other conditions related to muscular imbalance) may be concomitants of being in relatively unrecognized frightened and angry states with respect to the environment and therefore are to be interpreted as symptoms of these conditions rather than as causal factors in reading difficulties.

It is also not unlikely that, depending upon the strength of the fear-producing stimuli and the physical and genetic factors present:

*a*. Cerebral dominance is not able to be established in some instances giving rise to problems of handedness.

*b*. Speech centers are inhibited which result in speech disorders from mild stammering to complete aphasia, conditions which are also believed to be concomitants of being in a frightened state and which are not uncommonly associated with reading problems.

Finally, if reading is agreed to be a function of perception and learning these foregoing hypotheses with respect to the effects of fear upon reading may be expected to apply to processes of perception and learning in other areas with reading but an example of considerable cultural importance. Our general hypothesis is that while fear producing stimuli are of the highest valency in inhibiting many forms of conduct and behavior, including physiological as well as psychological processes, its opposite or courage-producing stimuli have the opposite effect and should perhaps receive increased emphasis in the culture and in psychological literature.

## REFERENCES

1. AMMONS, R. B., & HIERONYMOUS, A. N. Critical evaluation of a college program for reading improvement. *J. Educ. Psychol.*, 1947, 449-468.

2. BARBER, L. K. Immature ego development as a factor in retarded ability to read. *Dissertation Abstr.*, University microfilms, Ann Arbor, Mich., Publ. No. 3711, 1952, **12**, 503.

3. BEAR, R. M. Reading comprehension: Coöperative English test: lower and higher levels, C1 and C2. *In* O. K. Buros (*Ed.*), *Third Mental Measurements Year'b.* New Brunswick, N. J.: Rutgers Univ. Press, 1949.

4. BURFIELD, M. Identifying characteristics that retard reading in grades eleven to fourteen. In H. Robinson (*Ed.*), *Suppl. Educ. Monog.*, 1950, No. 72.

5. CANNON, W. B. The Way of an Investigator, a Scientist's Experiences in Medical Research. New York: Norton, 1945.

6. DAVIS, F. B. Iowa Silent Reading Tests (Rev.). *In* O. K. Buros (*Ed.*), *Third Mental Measurements Year'b.* New Brunswick, N. J.: Rutgers Univ. Press, 1949.

7. DUNBAR, F. (*Ed.*). Special sense organs exclusive of skin. In *Emotions and Bodily Changes.* New York: Columbia Univ. Press, 1946.

8. EBERL, M. Visual training and reading. *In* H. Robinson (*Ed.*), *Suppl. Educ. Monog.*, 1953, No. 77.

9. EPHRON, B. Emotional Difficulties in Reading. New York: Julian, 1953.

10. FERNALD, G. Remedial Techniques in Basic School Subjects. New York: McGraw-Hill, 1943.

11. GESELL, A. Vision and reading from the standpoint of child development. In H. Robinson (*Ed.*), *Suppl. Educ. Monog.*, 1953, No. 77.

12. IMUS, H. A., ROTHNEY, J. W. M., & BEAR, R. M. An Evaluation of Visual Factors in Reading. Dartmouth College Publ., 1938.

13. PEARSON, G. H. J. Survey of learning difficulties in the child. In *The Psychoanalytic Study of the Child*, Vol. VII. New York: International Univ. Press, 1945.

14. PFISTER, O. Love in Children. New York: Allen & Unwin, 1924.

15. ROBINSON, H. Why Pupils Fail in Reading. Chicago: Univ. Chicago Press, 1946.

16. ROBINSON, H. (*Ed.*). Emotional problems exhibited by poor readers. *Suppl. Educ. Monog.*, 1949, No. 68.

17. ROBINSON, H., & HUELSMAN, C. B., JR. Visual efficiency and progress in learning to read. *Suppl. Educ. Monog.*, 1953, No. 77.

18. ROBINSON, H. (*Ed.*). Treatment of poor readers with vision problems. *Suppl. Educ. Monog.*, 1953, No. 77.

19. SIEGEL, M. The personality structure of children with reading difficulties as compared with children presenting other clinical problems. *Microfilm Abstr.*, University microfilms, Ann Arbor, Mich., Publ. No. 2782, 1951, **11**, 1100-1101.

20. SIMPSON, E. Science Research Associates Reading Books 2 and 3. Chicago: Science Res. Associates, 1950.

21. STROUD, J. B. Reading comprehension: Coöperative English test: Lower and higher levels, C1 and C2. *In* O. K. Buros (*Ed.*) *Third Mental Measurements Year'b.* New Brunswick, N. J.: Rutgers Univ. Press, 1949.

22. TURNBULL, W. W. Iowa Silent Reading Tests. (Rev.) *In* O. K. Buros (*Ed.*), *Third Mental Measurements Year'b.* New Brunswick, N. J.: Rutgers Univ. Press, 1949.

23. WITTY, P., STOLARZ, T., & COOPER, W. Some results of a remedial reading program for college students. *Sch. & Soc.*, 1952, **76**, 376-380.

*Reading Improvement Program*
*Tulane University*
*New Orleans 18, Louisiana*

*The Journal of Genetic Psychology*, 1957, **90**, 203-212.

# THE EFFECT OF FRONTAL CORTICAL LESIONS ON ACTIVITY WHEEL AND OPEN-FIELD BEHAVIOR*

*Washington University School of Medicine*

John A. Stern[1]

## A. Problem

In a previous study (10) we reported on the effect of a series of electroconvulsive seizures on open field and activity wheel behavior. The present report deals with the effect of frontal area lesions on the same measures of behavior. The evidence, at the rat level, for changes in behavior attributable specifically to frontal area ablation is equivocal. Richter and Hawkes (8) report extensive changes in behavior of rats subjected to removal of the frontal poles. Their animals are described as savage, highly distractible, and showing perseverative behavior. Beach (1), who removed either the anterior neopallium or all brain tissue anterior to the corpus callosum, found little change in behavior of these animals as compared to control animals. He did, however, report that these rats showed more variability in behavior than did control animals, and found that in an enclosed field his anterior lesion animals more consistently followed the perimeter of the field, avoiding the center of the field more than unoperated or posterior area lesion animals did. Nauta and Brady (6) produced septal area injury in a group of rats and showed that this operation produced striking changes in behavior which proved to be of a transitory nature. During a period of a few hours after production of lesion these animals showed frenzied running to the slightest stimulation, attacking whatever was in their immediate vicinity.

The results of anterior cortical lesions on activity wheel running is less equivocal. Richter and Hawkes (8), Beach (2), and Lashley (3) all report that after frontal area ablation the rat shows an increase in activity as measured in activity wheels. Fulton (2), in a recent review, tempers this observation by commenting that, ". . . other animals supposedly with the same lesion failed to show this conspicuous restlessness" (p. 72).

In the present experiment open field behavior was evaluated pre-operatively as well as 30, 65, and 100 days after lesion. Activity wheel running was measured pre-operatively and 20 and 55 days after operation.

---

*Received in the Editorial Office on January 18, 1956.

[1] This experiment was conducted in the Psychological Laboratory of the University of Illinois.

## B. METHOD

### 1. *Subjects*

Subjects in this experiment were a group of 50 male albino rats reared in the animal laboratory of the Department of Psychology, bred from stock maintained in this laboratory. Animals in this experiment were weaned at approximately 25 days, were separated by sex, and had their ears marked. Animals were housed in groups of approximately 12 and always had a more than adequate supply of food (Purina chow) and water available. Age of animals at inception of experiment ranged between 100-115 days. The animals in this experiment were divided into three groups, 18 animals constituted the frontal lesion group, 16 a posterior area lesion group, and 16 animals in the control group. During the 155 days of the experiment 4 of the frontal lesion, 3 central lesion, and 5 control animals died. The results of this experiment are thus based on 14 animals in the frontal lesion group, 13 in the posterior lesion, and 11 in the control group.

### 2. *Apparatus*

The apparatus used for both measurement of open field and activity wheel behavior were described in a previous report (10).

### 3. *Procedure*

All animals were handled daily from a period 20 days prior to any experimental procedure through the 111th day of the experiment; thereafter they were handled routinely every fifth day as well as during any experimental procedure. Daily handling consisted of removal of animals from home cage, weighing them individually, and returning them to their home cage.

Measurement of activity wheel and open field behavior were previously (10) described and consisted of 10 daily measures of activity wheel running and five daily measures of open field activity.

### 4. *Surgical Procedure*

Cortical lesions and mock lesions were produced under aseptic conditions. The animal was placed under deep ether anesthesia and the hair from a level somewhat behind the ears to the level of the anterior aspect of the eyes clipped closely with scissors. The skull was then laid bare by a midline incision extending from the level of the ears to the posterior aspect of the eyes. Underlying tissue was removed by use of a high frequency cautery (power setting about 2). The skull was then trephined at the desired location by the use of a No. 4 trephine, two such holes being made, one on each side of the central suture. The anterior edge of the hole for the frontal lesion animals came to the level of the posterior aspect of the eye, the posterior edge being one to two mm. in front of the coronal suture. The trephine holes were separated from each other by approximately one to three mm. For the posterior area lesions trephine holes were located about two to four mm. behind the coronal suture, the two holes being approximately four mm. apart.

The cortical tissue underlying the burr hole and tissue anterior to, temporal to, and tissue lying between the burr holes was cauterized (power setting 5.5). An attempt was made to remove as much of the frontal cortical tissue as possible without damaging major blood vessels.

An equal amount of cortical tissue removal was attempted in the control lesions. After completion of brain removal, the opening was closed by suturing the skin flap with metal suture clips, four to five clips adequately closing the opening. The suture area was covered with collodion and the animal was returned to an individual living cage while still under the influence of the anesthetic agent.

All lesions were produced on the same day, time of production of lesion varying between the hours of 8:00 A.M. and 12:00 P.M. At termination of the experiment all lesion animals were killed and their brains were evaluated for extent of lesion. Amount of cortical tissue removed varied between 8 and 20 per cent, both lesion groups losing approximately equal amounts of cortical matter.

## C. RESULTS

We had hypothesized that frontal lesion animals would show changes in behavior similar to those shown by animals subjected to a series of *ECS*, while posterior area lesion animals would not show such changes. The specific hypotheses tested were that: (*a*) There would be no significant difference in any of our measures on the "pre-lesion" sample; (*b*) that the frontal lesion group would, in the open field, show a significant decrease in movement, increase in crouching, sum of "emotional indication," and in amount of defecation following operation and that these changes in behavior would be relatively enduring; (*c*) that the posterior lesion group would not show these changes, i.e., would behave similarly to a control group; (*d*) that the frontal lesion animals would show a significant increase in activity wheel running after operation.

The Mann-Whitney "*U*" Test (5) was utilized in treatment of data. One- and two-tailed tests of significance were utilized as indicated, one-tailed tests being utilized where specific changes had been hypothesized. Two types of analysis were used: (*a*) An analysis of "total scores"; with this measure comparisons could be made between experimental and control groups on the "pre-operate" and the "post-operate" measures. (For open field behavior the "pre-operate" measure of open field activity was initiated 10 days be-

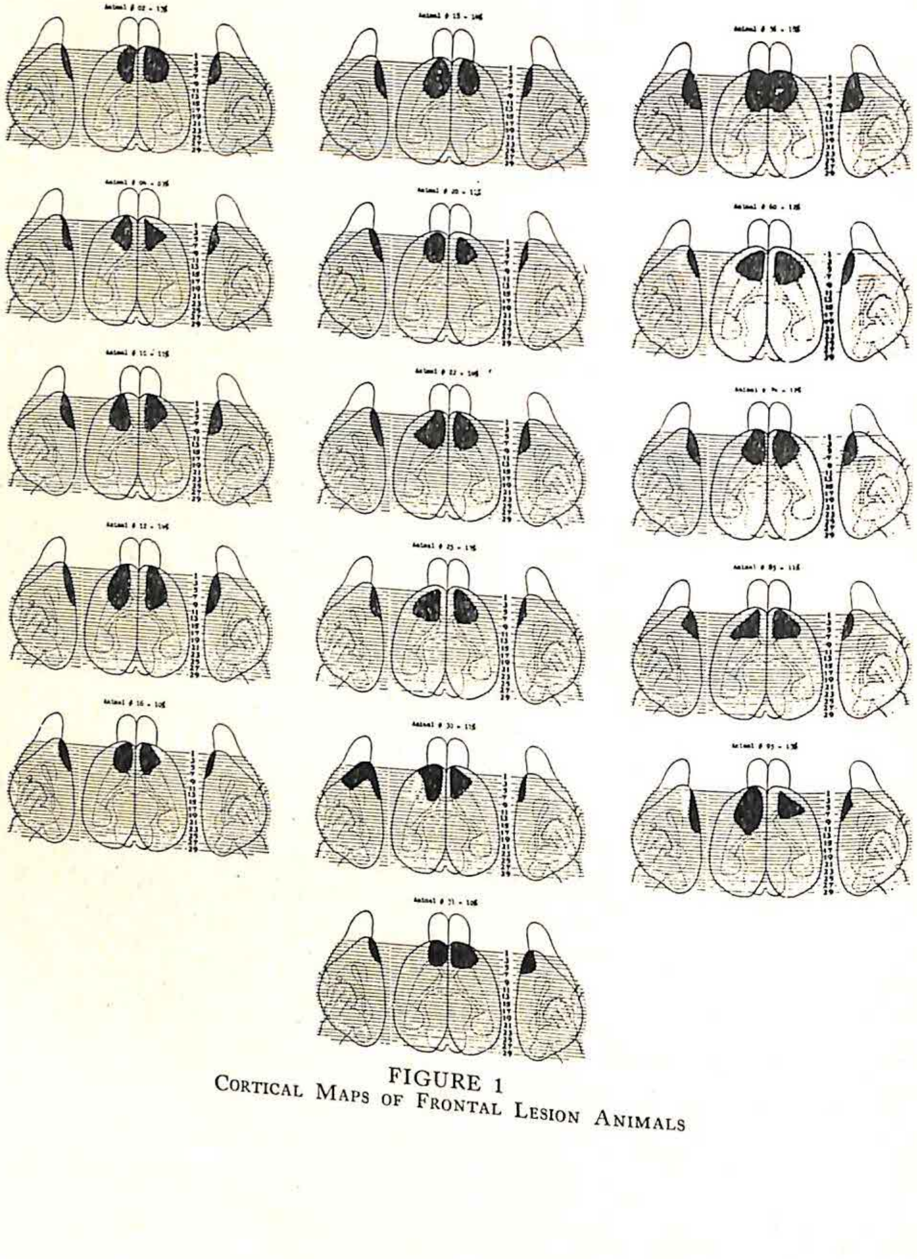

FIGURE 1
CORTICAL MAPS OF FRONTAL LESION ANIMALS

fore operation, the "pre-operate" activity wheel measures were initiated 20 days prior to operation. Post-operate measures for activity wheel running were initiated 20 days and 55 days after operation, those for open field activity were initiated 30, 65, and 100 days after operation. These samples will hereafter be referred to as the "pre" 1st, 2nd, and 3rd "post-operate" sample.) (*b*) The second type of analysis involved an analysis of "difference scores" based on the algebraic difference, for each animal, between his

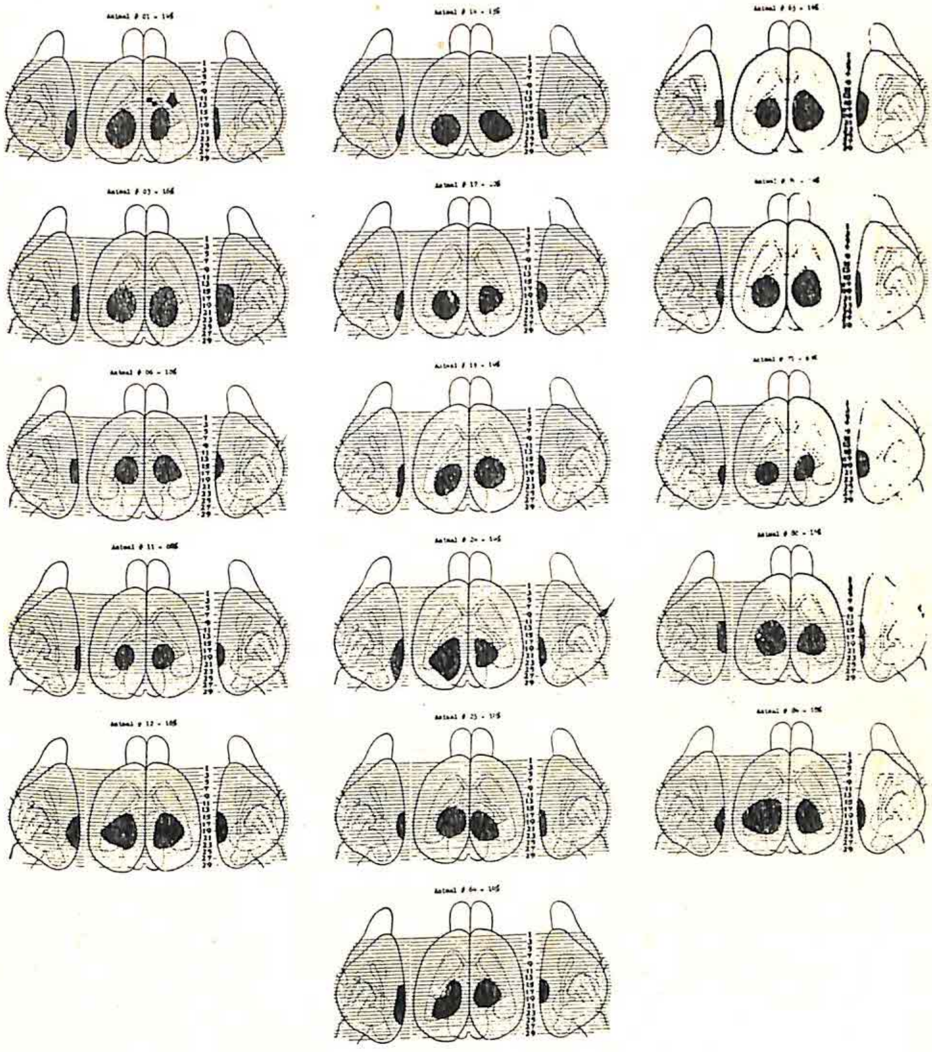

FIGURE 2
Cortical Maps of Posterior Lesion Animals

performance on any two samples of behavior. The results of both analyses reflect the same changes. The latter analysis was preferred since it shed light on the question as to where (between which samples) changes in behavior between groups compared were being manifested. Data for both open field and activity wheel running were analyzed in this manner. The results for the latter type of analysis are presented in Tables 1, 2, and 3.

TABLE 1

COMPARISON OF DIFFERENCE SCORES BETWEEN FRONTAL LESION AND CONTROL ANIMALS

| | Pre-1st post | Pre-2nd post | Pre-3rd post | 1st-2nd post | 1st-3rd post |
|---|---|---|---|---|---|
| Movement | .113 | .001 | .330 | .136* | .440* |
| Defecation | .121 | .004 | .163 | .064* | .322* |
| Crouching | .037 | .0002 | .0006 | .170* | .054* |
| Sum of emotional indicators | .019 | .0001 | .013 | .030* | .284* |
| Activity wheel running | .83* | .006* | | .20* | |

*Indicates *p* based on 2-tailed test of significance. All other *p*'s are based on 1-tailed tests.

In Table 1 a comparison of four measures of table top behavior as well as activity wheel running between frontal lesion and control animals shows that of the 14 comparisons made between the pre- and post-lesion measures 9 are significant beyond the .05 level. The measures which differentiated best appear to be "crouching" and "sum of emotional indicators." The comparison of difference scores between the posterior lesion and control groups shows that four of the 14 comparisons made between the pre- and post-lesion measures are significant at or beyond the .05 level. All four of these significant differences were found between the pre- and 2nd post-lesion set of measures. This change in behavior appears to be a transitory one, occurring after the first post-lesion testing and is no longer evidenced by the third post-lesion testing.

TABLE 2

COMPARISON OF DIFFERENCE SCORES BETWEEN POSTERIOR LESION AND CONTROL ANIMALS

| | Pre-1st post | Pre-2nd post | Pre-3rd post | 1st-2nd post | 1st-3rd post |
|---|---|---|---|---|---|
| Movement | .904* | .0004* | .824* | .0002* | .372* |
| Defecation | .780* | .250* | .540* | .262* | .952* |
| Crouching | .424* | .048* | .172* | .001* | .012* |
| Sum of emotional indicators | .872* | .002* | .164* | .001* | .322* |
| Activity wheel running | .31* | .018* | | .19* | |

*Indicates *p* based on 2-tailed test of significance. All other *p*'s are based on 1-tailed tests.

In Table 3 a comparison of difference scores between the frontal lesion and the posterior (control) lesion group similarly indicates the unexpected change in behavior on the part of the posterior lesion group on the second post-lesion testing. The change in behavior shown by this group is in the same direction as the changes noted for the frontal lesion group.

TABLE 3

COMPARISON OF DIFFERENCE SCORES BETWEEN FRONTAL AND POSTERIOR LESION ANIMALS

| | Pre-1st post | Pre-2nd post | Pre-3rd post | 1st-2nd post | 1st-3rd post |
|---|---|---|---|---|---|
| Movement | .020 | .834 | .397 | .012* | .076* |
| Defecation | .201 | .068 | .057 | .558* | .322* |
| Crouching | .0006 | .095 | .030 | .038* | .368* |
| Sum of emotional indicators | .048 | .268 | .032 | .204* | .864* |
| Activity wheel running | .70* | .46* | | .85* | |

*Indicates $p$ based on 2-tailed test of significance. All other $p$'s are based on 1-tailed tests.

## D. DISCUSSION

We had hypothesized that the frontal lesion animals would show a significant increase in activity when compared to control animals. Our findings indicate that there was no significant difference in activity wheel running between experimental and control animals on the first post-lesion measure. On the second measure a significant difference was found, but not for the reason hypothesized. It is a well documented fact that animals will show a decrease in activity with increasing age (7, 9). In our comparison we found that the frontal lesion animals showed less of a decrease in activity between the pre- and 2nd post-measure than was found for the control group, thus our hypothesis that frontal lesion animals should show an *increase* in activity as measured in activity wheels is not borne out. The posterior area lesion group showed a similar change in activity wheel running on the comparison between the pre- and 2nd post-measure. A number of authors (1, 3, 8) have found a significant increase in activity wheel running in animals subjected to frontal area lesions. Two possible reasons for our negative results suggest themselves; namely, that activity measures in our study were taken for too short intervals (10 daily 30-minute measures in the wheels) while in all of the above studies animals were either housed in or had free access to the activity drums for long periods of time. Our findings thus would suggest that short period measures are not sensitive enough to measure changes in activity for frontal lesion animals. For example, it is possible that these animals show spurts of activity limited to periods during which we were not measuring activity. A second possibility, and one more consonant with our thinking, is that frontal lesions produced in our experiment did not include the cortical area to which increases in activity (albeit not in the rat) have lately been referred to, namely ablation of the orbital surface of the frontal lobe, the head of the caudate nucleus, or severing the connections between these two areas. Cortical maps of the extent of lesions produced in our frontal area animals show them to be confined to the central aspects of the frontal cortex. It is thus a distinct possibility that our ablations did not incorporate those cortical areas to which increased activity has been attributed.

In a previous study (10) we showed that the effect of a series of 15 *ECS* administered at the rate of one per day produced long lasting changes in open field behavior. On the basis of theoretical speculations that both a series of *ECS* and frontal area lesions interfere with normal functioning of the thalamic and hypothalamic areas we expected that both procedures would produce similar behavioral changes. Our results indicate that such is apparently the case. Both frontal area lesions and a series of *ECS* produce similar and long lasting changes in behavior as measured in the open field. A series of 15 *ECS*, however, produced more severe changes than did our frontal area lesions. In both cases, however, an increase in "emotional" behavior was found. Another difference was that the *ECS* animals evidenced these changes in behavior sooner after experimental manipulation than was true for the frontal lesion animals. The former group showed all their changes at the first post-test measure while the frontal lesion animals did not show their maximal change during the 1st post-test. Changes in behavior developed somewhat more gradually than was true for *ECS* animals.

Our posterior lesion animals behaved similarly to the control group on both the first and third post-test, while they showed changes in behavior similar to that of the frontal lesion group during the 2nd post-test. These changes were not of the same order of severity as was found for the frontal lesion group. Another difference to be noted is that the change in behavior for the posterior lesion animals took place during the 2nd and 3rd post-test while those for the frontal lesion group proved to be more gradual (comparison of difference scores between 2nd and 3rd post-test for the frontal lesion and controls showed one out of five significance tests to be significant beyond the .05 level while in the comparison between posterior lesions and controls three of the five tests proved to be significant).

The finding of a rather sudden change in behavior on the part of posterior lesion animals as measured 65 days after operation is puzzling. Lashley

(3, 4) discusses a number of factors which might be invoked were it not for the puzzling fact that it took between 35 to 65 days for these changes in open field behavior to appear; i.e., no changes were noted at the first post-measurement conducted 35-40 days after operation. Since all degenerative changes are reported to be completed within three weeks after lesion, theories involving such a concept are thus eliminated. The author is not aware of any theories which would explain behavioral changes after cortical ablation where more than three weeks have elapsed between time of operation and appearance of behavioral change.

## E. Summary

The effect of frontal cortical lesions on a measure of activity wheel running and behavior in an open field, was compared to that of non-operate and posterior area lesion. The effect of the operative procedures on open field behavior were evaluated 30, 65, and 100 days after operation. Frontal lesion animals showed significant changes in behavior on post-lesion tests. These changes involved a decrease in activity in the open field and an increase in "emotional" indicators. The latter changes were most marked, persisting throughout the experiment. Activity wheel running showed that frontal lesion animals did not reduce activity (as a function of age) as much as control animals did. Posterior lesion animals showed a significant decrease in activity after lesion, as well as unexpected and transitory changes in open field behavior. These changes manifested themselves only on the 2nd post-test and were similar to the changes in behavior observed in frontal lesion animals.

## References

1. Beach, F. A. Effect of brain lesions upon running activity in the male rat. *J. Comp. Psychol.*, 1941, **31**, 145-179.
2. Fulton, J. F. Frontal Lobotomy and Affective Behavior. New York: Norton, 1951.
3. Lashley, K. S. Brain Mechanism and Intelligence. Chicago: Univ. Chicago Press, 1929.
4. ———. Factors limiting recovery after central nervous lesions. *J. Nerv. & Ment. Dis.*, 1938, **88**, 733-755.
5. Moses, L. E. Non-parametric statistics for psychological research. *Psychol. Bull.*, 1952, **49**, 122-143.
6. Nauta, W., & Brady, J. An exploratory study of the relationship between sub-cortical lesions and emotional behavior in the albino rat: The septal area of the forebrain. *Amer. Psychol.*, 1952, **7**, 257a.
7. Richter, C. P. A behavioristic study of the activity of the rat. *Comp. Psychol. Monog.*, 1922, **1**, 2.
8. Richter, C. P., & Hawkes, C. D. Increased spontaneous activity and food intake produced in the rat by removal of the frontal poles of the brain. *J. Neurol. & Psychiat.*, 1939, **2**, 231-242.

9. Slonaker, J. R. The normal activity of the albino rat from birth to natural death, its rate of growth and the duration of life. *J. Anim. Behav.*, 1912, **2**, 20-42.

10. Stern, J. A. The permanence of effect of a series of electroconvulsive seizures on open field behavior. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 411-415.

*Department of Psychiatry*
*Washington University School of Medicine*
*1420 Grattan*
*St. Louis 4, Missouri*

*The Journal of Genetic Psychology*, 1957, **90**, 213-218.

# HOARDING BEHAVIOR IN THE HOODED RAT AS A FUNCTION OF VARIED ALLEY ILLUMINATION*[1]

*Department of Psychology, University of Missouri*

MELVIN H. MARX, SHINKURO IWAHARA,[2] AND AARON J. BROWNSTEIN

## A. INTRODUCTION

Increased food hoarding in albino rats under conditions of high illumination of the alley has been reported by Stellar *et al.* (4). The present study was designed to test the influence of the alley illumination variable on the food hoarding behavior of the hooded rat. The fact that the eye of the *hooded* rat is normally pigmented would lead one to expect that the light variable might be of less importance than in the negatively photosensitive albino.

A secondary purpose of the study was to obtain observations on the hooded rat in order to determine whether it shows the same type of pre-hoarding behavior as has previously been observed in the albino (1, 2).

## B. METHOD

### 1. *Subjects*

*S*s were 19 female and 17 male hooded rats from the stock colony of the Department of Psychology at the University of Missouri. All animals were experimentally naïve. They were approximately 15 to 18 months old at the time of experimentation. An additional four male animals were discarded. Two developed ear disease, one died, and one failed to meet the time criterion during the preliminary phase.

### 2. *Apparatus*

The hoarding apparatus was that previously used and described by Marx (1, 3). The cages were 7 by 9 in. in floor area and 9 in. high. The hoarding alleys to which they were attached were 7 in. wide and 40 in. long. They were closed by 10-in. walls on the sides and ½-in. hardware cloth on the top. Metal bin trays were fitted into the far end of each of the six alleys in the apparatus.

---

*Received in the Editorial Office on January 19, 1956.

[1] This research was supported by a grant from the National Science Foundation.

[2] Now at Nara Women's University, Nara, Japan.

Illumination of the alleys was provided by a set of three electric light bulbs suspended directly over the bins at a height of two ft. For the high level of illumination 100-w. bulbs were used. For the low level 7.5-w. bulbs were used. A medium level of illumination was used only during the preliminary phase in order to avoid exposing the animals to either of the two experimental levels. It was provided by two 25-w. bulbs in the peripheral positions. Intensity of illumination at the home-cage entrance to the hoarding alleys, measured by a MacBeth Illuminometer, was approximately 1, 5, and 30 ml.

### 3. *Experimental Procedure*

*a. Accommodation phase.* All *S*s were placed in the individual living cages for 5 days. They were put on a 23-hr. feeding schedule and were given Purina Laboratory Chow pellets for one hr. each day.

*b. Preliminary phase.* Prior to feeding on the sixth day all *S*s were allowed access to the hoarding alleys for the first time. The food bins were in place, but empty. The medium level of illumination was used. Latency was recorded for first entrance into the alley, and for movement to a point halfway to the food bin. One hour following completion of this phase all *S*s were given a one-hr. feeding period.

On the basis of the latencies thus obtained the animals were divided into two major groups, *B*— and *D*—.

*c. Training phase.* Prior to feeding on the seventh day the *S*s were allowed access to the hoarding alleys. Ten food pellets were in each bin. Group *B*— animals were run under the high level of illumination (bright condition). Group *D*— animals were run under the low level of illumination (dim condition).

*S*s were allowed a maximum time of two hrs. to bring the 10 pellets into the cage. One-half hr. after the completion of training they were placed on *ad libitum* feeding with Purina pellets. This was maintained for the remainder of the experiment.

On the basis of the latency for the first pellet return and of the total time for return of all 10 pellets, Group *B*— and Group *D*— were each divided into two matched groups, according to whether the bright or dim illumination condition was to be present at test.

*d. Test phase.* On the eighth day *S*s were given access to the hoarding alleys with 100 pellets in each bin. Half of the *B*— animals were now tested under bright (Group *BB*) and half under dim (Group *BD*) illumination conditions. Half of the *D*— animals were similarly tested under bright (Group *DB*) and half under dim (Group *DD*) conditions. Five female and four male animals were in each group except *DB*, which had four females and five males.

A minimum time of 30 min. was allowed for the test period. If a *S* returned a pellet at any time after the first 10 min. the test was not concluded until he had gone for 20 min. without a return.

Detailed time records of all of the major movements of each *S* were made during test as well as training phase.

Four subsequent hoarding tests, each limited to a ½-hr. period without detailed observation, were then given.

*e. Reversal test phase.* The illumination conditions obtaining during test were then reversed for each group during a final series of five test days. They were run in the same manner as the test phase, except that the first reversal test was limited to 30 min.

## C. RESULTS[3]

### 1. *Matching of Groups*

The four groups were well matched with respect to performance measures in both preliminary and training phases. Differences among the groups in latency of first alley entrance and time to the half-way mark were nonsignificant ($Fs < 1.00$). Differences in time required to return 10 pellets in training were similarly nonsignificant.[4] These data were transformed to $\log_{10}$ to obtain homogeneity of variance.

### 2. *Test Results*

The results for the first test day are shown in Figure 1. Illumination conditions per se, on either training or test, did not make much difference. The most striking aspect of these results is the heightened hoarding activity of the two groups which had the same illumination conditions in training and test (*BB* and *DD*).

These conclusions are supported by the statistical analysis. Training and test variables each yielded $Fs < 1.00$, for first 30 min. or total test period.

---

[3] Summaries of the raw data for the material reported have been deposited as Document number 4805 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the Document number and by remitting $1.25 for photoprints, or $1.25 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

[4] Two of the *B*— animals did not return all of their ten pellets during training. They were given maximum time scores of 120 min. for purposes of this analysis. One was placed in Group *BB* and one in Group *BD*.

Training by test interactions were in each case significant ($F = 4.93$, $p < .05$, and $F = 5.97$, $p < .05$, for 1 and 32 $df$).

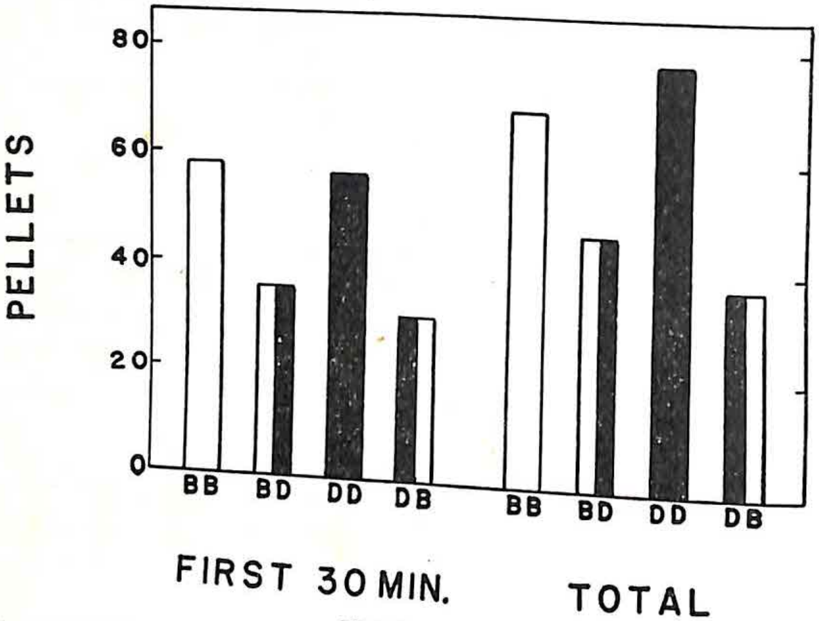

FIGURE 1
Mean Hoarding Scores for First Thirty Minutes and Total Time on First Test Day

The progressive changes in hoarding behavior among the four groups are shown in Figure 2. Diminishing group differences appeared as all groups showed extinction of the hoarding response.

No differences of any importance occurred during the reversal test, as can be seen in the right half of Figure 2.

### 3. *Comparison With Albino-Rat Behavior*

The 19 female hooded rats in the present study were compared with 40 female albino rats from another study which also used high drive conditions during preliminary and training phases. In spite of the disparity in age, the albinos being approximately three to five months old as compared with the mature hooded rats, essentially the same pre-hoarding behavior was found in both cases. The hooded females made almost exactly as many preliminary-phase alley responses before reaching the half-way mark (mean of 5.6, as against 5.8 for albinos). Alley entrances in training prior to the first pellet return were also almost identical (mean of .68, as against albino mean of .65).

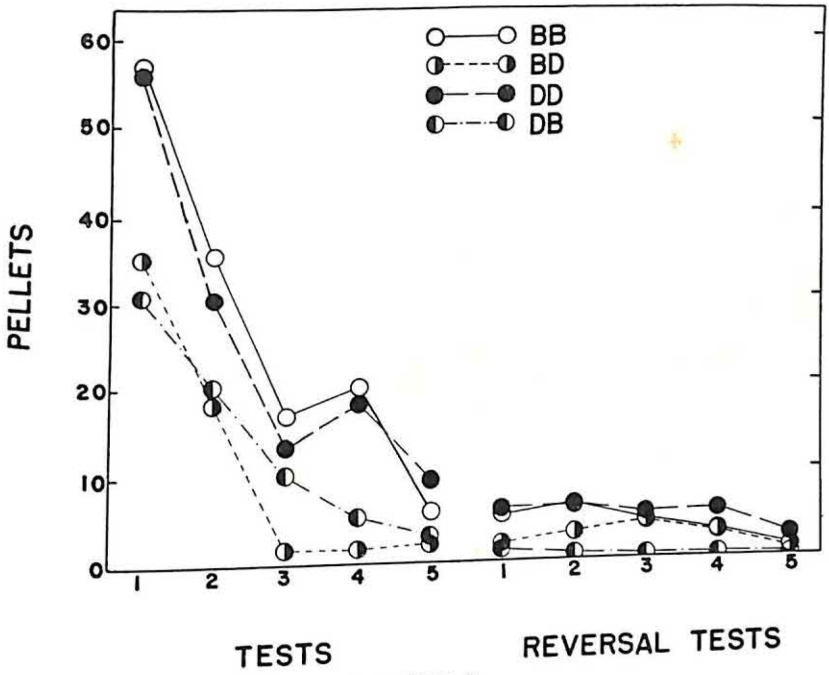

FIGURE 2
MEAN HOARDING SCORES FOR SUCCESSIVE THIRTY-MINUTE TEST AND REVERSAL-TEST PERIODS

## D. DISCUSSION

The major conclusion to be drawn from the present experiment is that no evidence has been found for hoarding behavior in the hooded rat to vary as a function of illumination conditions. Differences in illumination during training resulted in no significant differences in hoarding behavior in either training or test. Differences in illumination during test also resulted in no significant differences. The results of Stellar *et al.* (4) for the albino rat were not duplicated. The suggestion is that the hooded animal, with its markedly superior eye pigmentation, is not so immediately affected by gross visual (illumination) conditions.

The learning interpretation of hoarding (2) is supported by the significant reduction in hoarding scores following changes in illumination from training to test, since this points up the dependence of the habit upon the specific stimuli present during learning.

## E. SUMMARY

Two groups of mature hooded rats were trained to return food pellets with either bright or dim alley illumination. A series of five hoarding tests, with half of each group tested with bright and half with dim illumination conditions reversed from the earlier test, were given.

There were no significant differences between hoarding in bright illumination and hoarding in dim illumination. However, hoarding under illumination conditions differing from training conditions significantly decreased hoarding. The later change in illumination conditions had no effect on subsequent hoarding scores.

The hooded rats showed the same general pattern of pre-hoarding behavior as had previously been observed in the albino rat.

It is concluded that the hooded rat, with its normal eye pigmentation, is not so sensitive as the albino rat to gross illumination conditions per se, but does give evidence of learning the hoarding response.

## REFERENCES

1. MARX, M. H. Experimental analysis of the hoarding habit in the rat: I. Preliminary observations. *J. Comp. & Physiol. Psychol.*, 1950, **43**, 295-308.
2. ———. A stimulus-response analysis of the hoarding habit in the rat. *Psychol. Rev.*, 1950, **57**, 80-93.
3. ———. Experimental analysis of the hoarding habit in the rat: II. Terminal reinforcement. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 168-177.
4. STELLAR, E., HUNT, J. MCV., SCHOLSBERG, H., & SOLOMON, R. The effect of illumination on hoarding behavior. *J. Comp. & Physiol. Psychol.*, 1952, **45**, 504-507.

*Department of Psychology*
*University of Missouri*
*Columbia, Missouri*

*The Journal of Genetic Psychology*, 1957, **90**, 219-226.

# SPREAD OF EFFECT: ANIMAL STUDIES*

*Department of Psychology, University of Missouri*

MELVIN H. MARX

## A. INTRODUCTION

The relatively few studies concerned with spread of effect in infra-human animals are critically reviewed in the present paper, which may be considered as an addendum to the more extensive review of the spread-of-effect literature involving human subjects (8). The animal studies are reviewed separately from the human studies because to a large extent they have been conducted under basically different experimental procedures and involve different theoretical problems. In general they have been more closely related to some of the typical questions of basic learning theory.

Perhaps the most important procedural difference between the animal and the human studies is that the former have typically involved the "retained" situation (cf. 13). This means that each stimulus situation remains until the correct response occurs, as in the ordinary maze, rather than "vanishing," as in the characteristic Thorndikian multiple-choice serial learning design. Another major difference is that the stimuli and responses in the animal studies have typically been of the same sort from one situation to the next. A third important but unfortunate, because unnecessary, difference is that most of the animal studies have utilized the gross measure of total errors, or learning efficiency, rather than the more analytic one of tendency toward repetition of particular errors (cf. 8). As a result of these differences the artifacts that have been suggested as explanations of some of the positive gradient results obtained with animals have been of a different sort from those which need to be eliminated in human experimentation.

## B. THE BIDIRECTIONAL GOAL-GRADIENT HYPOTHESIS

### 1. *The Muenzinger, Dove, and Bernstone Study* (9).

The notion that the goal gradient might represent a special case of the Thorndike effect led these investigators to make this first animal study. They speculated that the typically observed backward gradient of errors might be only part of the story and that if appropriate measures were taken

---

*Received in the Editorial Office on January 19, 1956.

a similar gradient of error production extending forward from the goal might also be found.

The essential part of the study involved the use of a so-called "endless maze," that is, one which enabled the animal to proceed from one goal to the next without interruption, in order that the influence of the goal might be measured in a forward as well as a backward direction. This maze consisted of four straightaways set in the form of a square, with a goal box containing food at each corner and five choice-points between each pair of goal boxes. Four doors, one of which was unlocked, were present at each choice point. Experimental (goal-to-goal) animals were given a total of six uninterrupted circuits in the maze, thus providing 24 sets of goal-to-goal responses (data from all four straightaways being pooled in the analysis). Another set of 27 control (start-to-goal) animals were given 24 ordinary spaced runs (one a day) on a single straightaway. This spacing procedure was presumably introduced in the control group to avoid any possibility of a forward gradient being produced through temporal contiguity but it obviously makes direct comparison with the experimental subjects more tenuous.

The results were as predicted. The usual gradient of increasing errors with increasing distance backward from the goal was found in the control group. A clear-cut inverted *U*-type gradient occurred in the experimentals, with the choice point immediately following the goal showing practically as few errors as the point immediately preceding the goal. The difference between groups at the first choice point following the goal is said to have been highly significant (*CR* of 4.1). Moreover, this tendency to show fewer errors immediately following the goal was characteristic of most of the animals in the experimental group (81.5 per cent) but of only a small proportion of the control animals (33.3 per cent).

These investigators reject the goal-gradient hypothesis of Hull as an explanation of the experimental results on the ground that it rests on a goal-directed drive, and use a similar argument to dispose of Washburn's original explanation of backward elimination of maze errors. They then suggest that their bidirectional gradient is a special case of the Thorndike effect, noting that "There is more here than a superficial resemblance between the error curve . . . and a graph representing the repetition of wrong responses contiguous to a right one in Thorndike's experiment" (9, p. 237). This argument rests mainly upon the premise that "the excitatory tendencies for maze responses decrease forward as well as backward from a goal" (9, p. 238). The only direct evidence that is offered for such a basic resemblance, however, is the finding by Dove that "the uniformity with which the same error occurs at the same choice point in two successive trials of the goal-to-goal group is higher near the goals than midway between them" (9, p. 237). The paper to which reference is made for this has not been published, but the data are available in Dove's doctoral dissertation (1). If such a result were to occur, along with the presence of a much smaller total number of errors near the goals, this would present the peculiar situation in which animals tend to repeat a small number of specific errors near the goals while making a much larger number of different errors in the central choice points.

### 2. *The Hill Study* (4)

Using a single runway with four choice points and giving 72 rats five massed trials daily, Hill found only the usual backward wing of the goal gradient for the first 10 rewarded trials. Within the next set of 10 trials, however, the double or bidirectional gradient earlier found by Muenzinger, Dove, and Bernstone (9) became evident. Hill explains its appearance as a function of the relatively slow elimination in the two middle choice points of anticipatory errors, especially those that represent anticipations of the immediately succeeding correct choice. He thus rejects the interpretation in terms of a forward as well as a backward spread phenomenon. His own interpretation of these results is largely in terms of the assumed high degree of generalization among choice points.

### 3. *The Thompson and Dove Study* (12)

In an effort to verify the original Muenzinger, Dove, and Bernstone (9) results Thompson and Dove designed an improved version of a goal-to-goal maze and performed another experiment. They used a double-runway endless maze, consisting of two straightaways laid side by side. Each straightaway contained seven choice points, with seven doors per choice point, thus permitting a pattern in which each door was found open once in a given run. The two goal boxes were each opened to both straightaways, so that an animal would go down one straightaway and back the other, finding food reward in each goal box.

The bidirectional error gradient was again found. It appeared definitely in the group of 36 rats by the tenth trial and increased to a maximum by the twentieth run. A start-to-goal group of 51 rats, run as a control in a second experiment, gave the usual single-wing goal gradient.

The authors consider their results a confirmation of those earlier found by Muenzinger, Dove, and Bernstone. Their own theoretical interpretation is couched in terms of a backward and forward gradient spread from the goal of secondary reinforcing properties. They interpret the gradual appearance of the double gradient as an indication of its basic learning character. Moreover, they are vigorously critical of Hill's (4) interpretations of his results. They do not question the marked occurrence and slow elimination of anticipatory errors in the middle parts of Hill's straightaway but do question the explanatory value of his view, stating that "he fails to show the mechanism underlying the phenomenon" (12, p. 448). They also object to Hill's argument that the delayed appearance of the forward wing of the double gradient in his study is evidence against the spread hypothesis. The lack of strict comparability between procedures is pointed out, since Hill carried his rats from the single goal box to the starting box, thus not providing the same kind of goal-to-goal situation. Thompson and Dove suggest that the delayed appearance of the forward wing of the gradient may thus have been due to the fact that the animals needed to become emotionally adapted to this break in maze running before an effective goal-to-goal situation could develop.

This kind of disagreement over theoretical interpretation in terms of spread principles can hardly be clarified unless experimental measurements are made on the basis of the tendency to repeat particular errors, or correct responses, rather than merely on the basis of the gross number of errors observed. This requirement has been clearly met only in the studies by Jenkins (5, 6, 7) described below.

### 4. *The Haslerud Studies* (2, 3)

The two studies by Haslerud employed a rather complicated combination of free and forced choices in an elevated maze with diamond (two-choice) units.

In his first study (2) Haslerud was concerned with an "entrance gradient" as a counterpart to the usual goal gradient. This is to be differentiated from the bidirectional goal gradient since food was found only in the goal box and only a single trial per day was given, and in this paper the author does not refer to the earlier published research. Fourteen rats were given 20 trials on a 12-unit maze, with the first four turns opposite in direction to the last four and the middle four free. Some evidence is claimed for a short entrance gradient as well as a stronger, more extensive goal gradient, but the data are given only in graphic form and are not easily evaluated.

In his second study (3) Haslerud attempted to test the influence of a subgoal—food placed in the center of a 12-unit maze, supplementary to the feeding at the end—on the nearby responses. A general position preference was established in the 72 rats by forcing a choice at the first and last turns. The influence of forcing one or two turns to the opposite side at the point of supplementary feeding in the center of the maze was then investigated in several subgroups.

Briefly, the results gave evidence of a strong anticipation of the oppositely forced central turn, especially when it came immediately before the subgoal, and a weaker perseverative tendency. Although the author seems to find support for the proponents of the bidirectional goal-gradient hypothesis, and is critical of Hill's (4) interpretation in terms of anticipatory errors, this kind of reasoning is difficult to understand in view of the fairly apparent fact that Haslerud's data represent direct measurements of anticipatory and perseverative responses, with no attempt at measurement of response repetition, and therefore seem to contribute little to a separation of these two confounded tendencies.

## C. DIRECT TESTS OF THE SPREAD-OF-EFFECT HYPOTHESIS

### 1. *The Jenkins Studies* (5, 6, 7)

Jenkins has reported the first attempt to make a direct test with animals of the Thorndike effect, measuring the tendency to repeat free responses rather than errors as a function of distance from reward. He used a long elevated maze with as many as 20 diamond-shaped units, each providing a free two-choice situation at the bifurcation point.

In the first experiment (5) 35 rats were given 11 test trials with food reward on one of the short straightaways that connected the diamonds. A definite bidirectional gradient of decreasing response repetition was found around the point of reward. The response immediately preceding the straightaway on which the food was placed was repeated most frequently, but only slightly more so than the response immediately following. Repetition rate fell off regularly with increasing distance before and after the point of reward, with the exception of a slight inversion of the gradient at the fifth response following.

Although these results appear to offer strong evidence for the Thorndike effect one serious criticism has been raised. This is that, as Jenkins' own analysis demonstrated, the animals showed a strong continuing tendency to make the same turn as led to the reward both before and after the point of reward. A gradient of such anticipatory and perseverative errors also occurred and thus was confounded with the intertrial repetition-of-response gradient. Some suggestion that the Thorndike gradient was not entirely dependent upon such generalized responses was seen in the fact that it was strongest in the early trials whereas the anticipatory and perseverative tendencies did not reach their peak until the later trials. Nevertheless, these data cannot be accepted as unambiguous evidence of the Thorndike effect in this kind of experimental situation in the absence of a more convincing control.

Jenkins' other two published animal studies were essentially repetitions of the first one. They were conducted to measure the influence on the gradient of increased motivation (6) and increased incentive (7). In general the same results were found as in the first experiment. A somewhat higher rigidity of responses was found under higher drive, and doubling the amount of the incentive seemed to increase the level of repetition of responses before reward more than after.

## D. Extinction Tests in a Nonserial Situation

### 1. *The Nagaty Studies* (10, 11)

The two experiments recently reported by Nagaty represent a first attempt to test the spread-of-effect hypothesis within a Hullian type of theoretical framework. Conducted as doctoral research in the laboratory of N. E. Miller, they were apparently motivated by the inability of Hullian reinforcement theory to account for the strengthening of responses which follow rather than precede reinforcement.

Nagaty's basic design involved the conditioning of a specific response to a stimulus in the usual way—that is, with reinforcement following the response—and the subsequent testing of that same response under extinction conditions with the reinforcement now *preceding* the conditioned stimulus by a variable time interval. According to the author's interpretation of the spread-of-effect principle, continued increments in response strength rather than extinction should occur under these conditions, especially for the group with the short time interval.

The first experiment (10) utilized an avoidance-conditioning technique with buzzer as the *CS* and shock as the *UCS*. The 40 rats were trained to a fairly high level of avoidance-response strength (10 successive wheel-turning responses to buzzer alone within two seconds). They were then given 50 trials in which the shock preceded the buzzer by 1 sec., by 20 sec., or was omitted as in the normal extinction procedure. In an effort to control for any immediate artifactual influences resulting from the different testing procedures an additional 50 trials were given with one-third of each of the three major groups now tested under each condition.

Decrements in both amplitude and frequency of the wheel-turning response occurred for all three groups throughout the first 50 extinction trials, although the decrement in frequency for the 1-sec. group was not statistically significant. Significant differences were found in the rate of extinction, with greatest resistance in the 1-sec. group and least in the no-shock group. This result was in the direction that was predicted but was apparently not anticipated by the investigator. It is pointed out, in an ad hoc manner, that these differences "could easily have been produced by fear perseverating from the shock" (10, p. 243).

Correspondingly significant differences were not found in the last 50 test trials, although the frequency scores for the first 10 of these trials did show significant variation. The importance of these last 50 trials, however, is somewhat limited in view of the fact that most of the extinction had already occurred within the first set of 50 trials.

The second experiment (11) was conducted with 41 rats trained in a bar-pressing response with food reinforcement. The design was essentially the same. No important differences among groups occurred in extinction trials.

Several major points may be made in criticism of the design used in these studies. Spread-of-effect experimentation has been performed in situations involving a *series* of *different stimulus-response relations* with the critical error responses having a relatively *low order of strength*. Prolonged strengthening of a single response and the subsequent extinction of it under conditions such as Nagaty used must then be recognized as a most extreme test of "the authenticity of the claimed after-gradient of reinforcement in both 'spread of effect' and 'backward conditioning' phenomena" (10, p. 240). The relevance of the design to the backward-conditioning problem seems somewhat closer.

## E. SUMMARY AND CONCLUSIONS

This paper has critically reviewed the few studies concerned with the spread-of-effect phenomenon as evidenced in infra-human subjects.

Further experimentation with increasingly complex situations will be necessary before any clear-cut statements about the "authenticity" of the spread hypothesis can be made upon the basis of research with animal subjects. As a next step, for example, one might suggest a design utilizing as the critical test response one other than that involved in the reinforcement (a different manipulandum, for example, if Nagaty's general situation were retained). Accumulation of results from this kind of experimentation may be expected to indicate whether the spread-of-effect phenomenon, if validated with human subjects, can also somehow be applied in simplified instrumental animal learning.

In spite of the limitations in the research thus far reported, it is concluded that this kind of experimentation offers great promise of effective investigation of some of the basic problems involved. The surprisingly small amount of animal research thus far conducted on the spread phenomenon needs to be supplemented.

## REFERENCES

1. DOVE, C. C. The effect of successful and unsuccessful responses upon uniformity and variability of behavior. Unpublished doctor's dissertation, Univ. Colorado, 1936.
2. HASLERUD, G. M. Form and interaction of goal and entrance gradients in a unilinear maze. *J. of Psychol.*, 1945, **20**, 249-257.
3. ———. Properties of bi-directional gradients at subgoals. *J. Genet. Psychol.*, 1949, **74**, 3-16.
4. HILL, C. J. Goal gradient, anticipation, and perseveration in compound trial-and-error learning. *J. Exp. Psychol.*, 1939, **25**, 566-585.
5. JENKINS, W. O. Studies in the spread of effect: I. The bi-directional gradients in the performance of white rats on a linear maze. *J. Comp. Psychol.*, 1943, **35**, 51-56.
6. ———. Studies in the spread of effect: II. The effect of increased motivation upon the bi-directional gradient. *J. Comp. Psychol.*, 1943, **35**, 57-63.
7. ———. Studies in the spread of effect: III. The effect of increased incentive upon the bi-directional gradient. *J. Comp. Psychol.*, 1943, **35**, 65-72.
8. MARX, M. H. Spread of effect: a critical review. *Genet. Psychol. Monog.*, 1956, **53**, 119-186.
9. MUENZINGER, K. F., DOVE, C. C., & BERNSTONE, A. H. Serial learning: II. The bi-directional goal gradient in the endless maze. *J. Genet. Psychol.*, 1937, **50**, 229-241.
10. NAGATY, M. O. The effect of reinforcement on closely following *S-R* connections: I. The effect of a backward conditioning procedure on the extinction of conditioned avoidance. *J. Exp. Psychol.*, 1951, **42**, 239-246.
11. ———. The effect of reinforcement on closely following *S-R* connections: II. Effect of food reward immediately preceding performance of an instrumental conditioned response on extinction of that response. *J. Exp. Psychol.*, 1951, **42**, 333-340.
12. THOMPSON, M. E., & DOVE, C. C. A further study on the bi-directional goal gradient in the endless maze. *J. Exp. Psychol.*, 1943, **33**, 447-456.
13. WAITS, J. V. The law of effect in the retained situation. *Arch. of Psychol.*, 1937, **30**, No. 208. Pp. 56.

*Department of Psychology*
*University of Missouri*
*Columbia, Missouri*

*The Journal of Genetic Psychology*, 1957, **90**, 227-238.

# THE DISTINCTION BETWEEN LIVING AND NOT LIVING AMONG 7-10-YEAR-OLD CHILDREN, WITH SOME REMARKS CONCERNING THE SO-CALLED ANIMISM CONTROVERSY*

*Teachers College, Falun, Sweden*

---

GÖTE KLINGBERG

---

## A. THE PROBLEM

In the literature of child psychology animism is defined as a tendency to ascribe life and consciousness to all things. Thus the term is used in the same sense as ethnology uses the term animatism. According to Piaget, this animism is a characteristic trait of a primitive mental structure. In the life of children a first phase of animism, in which it is connected with activity in general, especially with activity of use for man, reaches up to 6-7 years. In the second phase of animism the child attributes life and consciousness to everything that moves. The mean age is 6-8 years. In the third phase, with the mean age between 8-9 and 11-12 years, only things that move by themselves are affected by the animistic tendency. Towards 11-12 years three-fourths of the children reach the fourth phase, in which life is attributed only to animals or to animals and plants (3, p. 159 seq.).

The discussion around Piaget's concept of animism is reported in an article by Huang in this journal of 1943 (1). *The Journal of Genetic Psychology* has also published several other important articles in the so-called animism controversy. Russell and Dennis tried to construct a standardized procedure for the investigation of animism, in which they asked the children about 20 objects. In a research on 774 American children, published by Russell in 1940, it was found that the universality of the stages of animism could not be disputed, though the age range of the stages could not be so distinctly limited as Piaget had attempted (4 and 5).

In an article from 1945 Huang (with Lee) reports a similar method (2). Forty Chinese children were asked about 10 objects: dog, tree, river, stone, pencil, bicycle, ball, automobile, watch, the moon. Huang and Lee asked not only if the object was living but also if it had life and five other questions concerning concepts assumed to be allied to animism: if the object could feel, etc. The children were between 3;5 and 8;7 years old. The

---

*Received in the Editorial Office on January 23, 1956.

result was that an object without life is said to be alive only sporadically. More correct answers were given to the question Has it life? than to the question Is it living? One cannot be sure of what the child exactly means by living. On the whole Huang is very critical about Piaget's results. One has not to assume that the children go from a general animistic tendency to a more physical concept of reality but that the whole problem in the beginning is very diffuse and that the distinction between living and not living is grasped gradually. The probability that an object would be mistaken as living increases to the extent to which the object has traits in common with living beings.

In 1951 Strauss published a re-examination of the Huang-Lee data in the same journal (7). His opinion is that the results of Huang and Lee could be interpreted in both ways. To the question Has it life? the children did not give many animistic answers. On the other hand, to the question Is it living? the animistic answers were numerous. Strauss agrees that not all inanimate objects are equally hard to classify correctly. It is also true that Piaget's studies dealt largely with the child's concepts of far-away and also for adults mystifying objects. However, one cannot say that Huang and Lee have shown that young children do not display a general tendency to endow inanimate objects with life and consciousness. The main problem has not yet been decisively solved.

From my point of view the problem in reality is two problems and it is therefore meaningless to answer it with yes or no. The two problems that must be dealt with separately are: (*a*) Do the children attribute life to not-living objects? (*b*) If so, is it to be assumed that the cause is an animistic tendency based on a primitive mental structure in Piaget's meaning?

## B. The Experiment

The first problem must be the easier to deal with empirically. In order to get material of our own my pupils at the Teachers College at Falun, Sweden, have interrogated 97 7-10-year-old children at the practising-school, especially 7-8-year-olds. Thirty-eight were interrogated in 1951-52 by Miss Marta Eriksson, the rest by teamwork in 1952-54. The method was that of Huang, in the beginning all of Huang's questions were used, later on only five of them.

The difficulty is to know what the child means when it says that the river, etc., lives. It is therefore naturally a good idea in itself to widen the interrogation to other questions such as Does it feel anything if I prick it with this needle? However, as Piaget (3, p. 200 seq.), Russell (6), and

Huang-Lee have shown, the correlation between the answers to different questions of this general type is not very great. I consider it methodically disputable to construct a general way of thinking called animism and to examine it with different questions. What exactly is possible to investigate are the answers to specific questions. The difficulty is clearly demonstrated in our new material. In Table 1 a comparison is made of the answers of the 7- and 8-year-olds to the first and second questions (Is it living? and Does it feel anything if I prick it with this needle?) concerning four objects. The answers of a very few of the children are omitted. These children have answered I don't know on one or both of the questions or their answers have been judged as unclear.

TABLE 1
ANSWERS OF THE 7-8-YEAR-OLDS CONCERNING FOUR OBJECTS TO THE QUESTIONS IS IT LIVING? AND DOES IT FEEL ANYTHING IF I PRICK IT WITH THIS NEEDLE?

| Object | *N* | Lives and feels | Lives but does not feel | Does not live but feels | Does not live and does not feel |
|---|---|---|---|---|---|
| Tree | 65 | 24 | 33 | 0 | 8 |
| River | 65 | 1 | 23 | 0 | 41 |
| The moon | 59 | 30 | 12 | 2 | 15 |
| Watch | 66 | 3 | 21 | 0 | 42 |

It is obvious from such a distribution of the answers that the two questions must be very different questions. What is to be inferred from the fact that scarcely anyone says that the river and the watch feel anything but not a few that they live? Possibly something about what the children mean by living. But the conclusion is not so easily drawn. In agreement with half of the children most adults would answer that the tree lives but that it does not feel. The task of determining, on the basis of five or seven answers, whether the child thinks in an animistic way or not, is often insoluble. As this article is intended to deal with the distinction between living and not living, I consider it right to limit the report to the answers to the first and last questions: Is it living? and Has it life? It cannot be avoided that the question what the child really means by living often will appear.

The results from these two questions to the 7-8-year-olds and to the 9-10-year-olds are shown in Table 2 and Table 3. The 7-8-year-olds number 67 (30 boys and 37 girls), the 9-10-year-olds 30 (16 boys and 14 girls).

The experimental situation is a bit bewildering, and the results are not absolutely valid in the individual cases. Still, all answers together give an impression of real tendencies. The different objects are described as living to a very varying extent. Almost all say that the dog lives but scarcely anybody that the stone lives. Thus the distinction between living and not living is made in the right way.

Some objects, however, which adults regard as not living, are stated to be living. That children attribute life to some not-living objects is in fact shown by Piaget, Russell and others and also to some extent by Huang and Lee. The new Swedish material gives the same answer and it concerns relatively old children. Thus the first of our questions is settled. The children attribute life to not-living objects. But, it must be added, not to all not-living objects.

The extent to which the different objects are described as living and as having life is shown in Table 4. Only the answers of the 7-8-year-olds are treated as being the most representative material and in addition the most adequate for a comparison with the Huang-Lee data, their children not being older than 8;7.

TABLE 4
THE EXTENT TO WHICH THE 10 OBJECTS ARE DESCRIBED AS LIVING AND AS HAVING LIFE (7-8 YEAR OLDS, $N = 67$)

| Question: Is it living? | | Question: Has it life? | |
|---|---|---|---|
| Object | Described as living, percentage of children | Object | Described as having life, percentage of children |
| Dog | 94 | Dog | 92,5 |
| Tree | 88 | Tree | 74,5 |
| The moon | 67 | The moon | 61 |
| River | 37,5 | River | 25,5 |
| Watch | 37,5 | Watch | 22,5 |
| Automobile | 33 | Automobile | 21 |
| Bicycle | 21 | Pencil | 13,5 |
| Ball | 16,5 | Ball | 10,5 |
| Pencil | 12 | Bicycle | 9 |
| Stone | 6 | Stone | 4,5 |

Huang and Lee give the following order for the animation of the not-living objects in regard to the question Is it living?: the moon, watch, bicycle, automobile, ball, river, stone, pencil, and in regard to the question Has it life?: the moon, watch, automobile, ball, bicycle, pencil and stone, river. The correspondence between the Chinese and Swedish materials is striking with the exception that the river to a larger extent is described as living by the Swedish children than by the Chinese.

Huang and Lee found that the question Has it life? gave more correct answers as to the not-living objects and less correct answers as to the tree than the question Is it living? Strauss asks whether this fact is a function of the Chinese language. It is therefore interesting to note that the same tendency is clearly recognizable in the Swedish material. (Besides in Table 4 this tendency is to be seen for the older children in Table 3.) The tendency thus seems to be universal. Perhaps the cause is that it is more likely that the word living is associated with other concepts as for instance moving and sounding than the word life, which is more abstract. This is only a suggestion.

## C. Discussion

It may be said, in accordance with Strauss' critical views, that Huang and Lee have stated their conclusion too categorically. However, and this is also the opinion of Strauss, the answers of the children are more correct than Piaget has alleged. In the second phase of animism, with the mean age of 6-8 years, the children, according to Piaget, attribute life to everything that moves. It will be seen from Table 2 that of the 7-8-year-olds in the Swedish material only 21 per cent say that the bicycle is living and 9 per cent that it has life. As to the ball the percentages are 16.5 and 10.5 resp., as to the automobile 33 and 21 resp. Objects which move by themselves are, according to Piaget, attributed with life by children with the mean age between 8-9 and 11-12 years. What objects move by themselves besides the objects which also adults regard as living? The watch perhaps, if one does not understand the mechanism. It is of interest that even 33 and 23 per cent resp. of the 9-10-year-olds attribute life to the watch. However, only 37.5 and 22.5 per cent resp. of the 7-8-year-olds do the same. The river no doubt moves by itself, if one does not have a good understanding about what a river is; 37.5 and 25.5 per cent resp. of the 7-8-year-olds say that it is living or has life, even 23 and 20 per cent resp. of the 9-10-year-olds say the same. Yet only one-third of the 7-8-year-old children give this answer. The moon moves by itself, if one does not know the laws of physics. It is interesting that 67 and 61 per cent resp. of the 7-8-year-olds and even 43 and 30 per cent resp. of the 9-10-year-olds say that it is living or has life. The moon is the only of the discussed not-living objects to which more than half of the 7-8-year-olds attribute life. But, one may ask, is movement a more characteristic feature of the moon than of the river? It seems more likely that fewer will observe the movement of the moon than of the river. Is there no other cause than the movement to the animation of the moon? Table 4 gives the not-living objects very much in the order in which their physical traits are familiar to the children. The growth of knowledge must be fundamental for the classifying of the surrounding objects into the categories of living and not living. Huang and Lee have stressed too much their view that the children generally do not regard not-living objects as living. Their main point, however, seems to me to be right. Rather than from a universal animism to physicalism, the growth goes from ignorance to better and better knowledge. The distinction between living and not living is difficult and can sometimes be hard even to adults. That more not-living objects are placed in the category of living things than living objects in the category of not-living things may naturally be due to the influence of such traits as moving and sounding. Intentional or unintentional suggestions from the environment also sometimes are likely to make the distinction difficult. As long as the children have not come into contact with astronomy they have very little real knowledge of the moon. But they have heard about the man in the moon and about how the moon peeps in through the window.

Strauss thinks that the Huang-Lee data can support both the hypotheses of Piaget and Huang. But the question whether children attribute life to some not-living objects or not has nothing to do with the main hypothesis of Piaget. The hypothesis of Piaget is that the so-called animistic tendency is based on a primitive mental structure. But that is our second question which must be dealt with separately. The existence of a general animistic tendency in very young children cannot in fact be shown empirically. Even in the first phase of animism described by Piaget the animistic tendency depends on an activity of the object in question. The real first phase in which everything is interpreted as living must come so early in childhood that it is practically impossible through interrogation to form an opinion of the experience of the child. Instead, this first phase of animism is based on theoretical considerations. It must not be overlooked that Piaget is influenced by certain theories which may be false. Thus he is influenced by the doctrine of primitive mentality above all founded by the French sociologist Lévy-Bruhl. He further accepts the "genetic parallels," based on the recapitulation theory. But the picture of the primitive thinking of primitive man is antiquated in the science of ethnology.

The second of our questions must therefore be answered in the negative. There are no findings that suggest that the thinking of children, even before the age of seven, is a function of a fundamentally other mental structure than that of the adult. However "animistic" the answers of children may seem they cannot demonstrate a primitive thinking in Piaget's meaning. This is not the same as to say that the children do not have any difficulty with the distinction between the concepts of living and not living. The distinction is no doubt early made and in school-age rightly applied in most cases. Mistakes are however common. These do not show the animism of Piaget but they show the difficulties of the children to come to a true understanding of the surrounding world. So they are of great interest for child psychology and pedagogics.

## D. SUMMARY

1. In the "animism controversy" the existence of child animism in Piaget's meaning is disputed. It is said that the problem has not yet been decisively solved. In reality, however, two problems must be dealt with separately: (*a*) Do the children attribute life to not-living objects? (*b*) If so, is it to be assumed that the cause is an animistic tendency based on a primitive mental structure in Piaget's meaning?

2. The first problem is dealt with in an investigation made at the Teachers College at Falun, Sweden, where 97 7-10-year-old children (67 7-8 and 30 9-10-year-olds) were interrogated with the method of Huang. The results are tabulated and discussed. It will seem that the distinction between living and not living is much better applied at this age than Piaget thought. In this point the criticism of Huang and others against Piaget is verified. On the other hand one must observe that even at the age of 9-10 years the distinction is difficult for some children in regard to certain objects. The criticism of Strauss against Huang is stressed in this respect.

3. The second question, whether the explanation of the difficulty to make distinctions between living and not-living objects is to be found in insufficient knowledge and suggestions from the environment or in a specific way of thinking belonging to a primitive mental structure, cannot be solved through such empirical investigations. In view of the fact that Piaget is influenced by the recapitulation theory and the doctrine of the primitive mentality of primitive man—theories that have little connection with modern science—the explanation of Piaget seems to have little or no probability. The actual difficulties to distinguish between living and not living in regard to certain objects in some children even at the age of 9-10 years are, however, of interest, and, because of the probability of many similar difficulties in regard to other distinctions and concepts, of practical significance in pedagogics.

## REFERENCES

1. HUANG, I. Children's conception of physical causality: A critical summary. *J. Genet. Psychol.*, 1943, **63**, 71-121.
2. HUANG, I., & LEE, H. W. Experimental analysis of child animism. *J. Genet. Psychol.*, 1945, **66**, 69-74.

3. PIAGET, J. La représentation du monde chez l'enfant. Paris: Alcan, 1926.
4. RUSSELL, R. W., & DENNIS, W. Studies in animism: I. A standardized procedure for the investigation of animism. *J. Genet. Psychol.*, 1939, **55**, 389-400.
5. RUSSELL, R. W. Studies in animism: II. The development of animism. *J. Genet. Psychol.*, 1940, **56**, 353-366.
6. ———. Studies in animism: IV. An investigation of concepts allied to animism. *J. Genet. Psychol.*, 1940, **57**, 83-91.
7. STRAUSS, A. L. The animism controversy: Re-examination of Huang-Lee data. *J. Genet. Psychol.*, 1951, **78**, 105-113.

*Teachers College (Folkskoleseminariet)*
*Falun, Sweden*

*The Journal of Genetic Psychology*, 1957, **90**, 239-253.

# ACADEMIC ACHIEVEMENT IN A GROUP OF MENTALLY DISTURBED ADOLESCENTS IN A RESIDENTIAL TREATMENT SETTING*

*Biometrics Research, New York State Psychiatric Institute*

Kurt Salzinger[1]

## A. Introduction

In recent years residential treatment has come to the fore as a new concept in the treatment of mentally disturbed adolescents. While both papers and books (1, 2, 3, 8, 9, 10) have appeared describing the methods and concepts, and infrequently some of the results of such treatment, few attempts have been made to examine any of these aspects by any but clinical or impressionistic methods. This study which was conducted at the Linden Hill School is part of a larger program of research whose purpose it is to analyze objectively some of the data concerning the children under treatment, the staff members who treat, the setting and policy of the institution, the theoretical basis of residential treatment as well as the effect of the treatment process upon the children. An attempt has been made to examine the academic achievement of mentally disturbed children who receive residential treatment.

Academic achievement is an important area of functioning that must be assessed in any mentally disturbed child if only because much of any child's life is spent in school and in learning new things. Abnormality in academic achievement may be looked upon both as a cause and a result of mental disturbance. When acting as a cause, then change in academic achievement indicates a change in the mental stress at least partially responsible for the disturbance. Where it is the result, the change may be viewed as a change in the mental disturbance itself.

In a way, this study is complementary to the series of papers written by Feinberg (4, 5, 6), since it has also undertaken to examine academic achievement in various areas in a particular set of exceptional children. While his studies deal with orphans in foster homes and in orphanages, as well as delinquent children in a special home for delinquents, this treats mentally

---

*Received in the Editorial Office on February 3, 1956.

[1] The author wishes to acknowledge the help of Dr. Leonard Kornberg, Principal of the Linden Hill School, and his teaching staff consisting of Sadie Moody, Mrs. Pat Esterson, and Meyer Sussman for making the test data available. This paper has profited also from some comments made by Dr. Kornberg, Marvin Scherer, and Mrs. Edith Smith. This study was done while the author was at the Linden Hill School.

disturbed non-delinquent children in a residential treatment setting. This study differs from those of Feinberg insofar as it attempts to relate achievement and especially change in achievement (including assessment of the change as improvement or deterioration) to other variables such as sex, age, *IQ*, diagnosis, original achievement, age adequacy of the original achievement and subjects taught to the children between tests (i.e., the specific effectiveness of teaching particular courses to such children). It might also be mentioned that this study is also concerned with examination of the usefulness of this type of data in studying disturbed children.

## B. METHOD

### 1. *Population*

The Linden Hill School presently houses a group of 26 mentally disturbed children, 13 girls and 13 boys. They range in age from 13 to 18 with a median age of 15 years, and in *IQ* from 68 to 127 with a median value of 99. The children's diagnoses are distributed in the following manner: schizophrenia or schizophrenic reaction—11, schizo-affective disorder—1, obsessive-compulsive disorder—2, personality deviation with effeminate and homosexual trends—1, infantile personality—3, infantile character structure—1, severe character neurosis—2, character disturbance—passive-aggressive type—1, deferred—4.

All of the children being treated at Linden Hill could not be included in this study for the following reasons: Two children had to be excluded because they are unable to read, and therefore could not take the tests at all; two children came too late for the first test administration; one child was hospitalized for somato-therapeutic treatment during the second test administration. In each analysis all of the available data was used.

Since all children were not exposed to the same subjects it was feasible to separate out four groups, i.e., those who studied English against those who did not, and those who studied arithmetic against those who did not. Eleven children were exposed to English classes, the rest of the Linden Hill children were not, 17 children were exposed to arithmetic while the rest were not. Only arithmetic and English class divisions will be considered for this study since these are the only two subjects which the achievement test tapped.

A note of caution must be injected concerning the subject groupings. To begin with, these groupings held only for the spring term of the year intervening between the tests. In the fall term the school program underwent complete reorganization with such constant shifting of the student body as to defy making up a consistent grouping. Another factor to be considered in evaluating the analysis of the subject grouping is the selection process which went into setting up different classes; academic ability, emotional readiness to learn, relationship to each particular teacher, relationship to other children, the limiting factor of only three teachers, etc., were the factors considered.

### 2. *Instrument and Procedure*

The Stanford Achievement Test consisting of paragraph meaning, word meaning, spelling, language, arithmetic reasoning, and arithmetic computation was administered in a standard manner according to the official directions (7). Different forms of the test were administered to the children, the first in September, 1954 (the beginning of the school year), and the second in June, 1955 (the end of the school year).

### 3. *Types of Scores*

Three types of scores were used for some analyses, two for others, and only one for the rest. Each of the types of scores is useful in its own way; one is the modal-age grade scores which are computed in terms of the modal age group in each grade and therefore leave out any exceptionally bright or dull children; since the latter outnumber the former, the total-group grade scores (the second type score) yield higher scores for the same achievement than the modal-age grade scores. The third type score used in this study was the *K* score which constitutes an equal interval scale; in other words, this type of score allows us to measure change in terms of actual achievement rather than in grade years. Grade scores may sometimes give us an erroneous picture, since change from grade to grade is not equal, e.g., the change in word meaning from Grade 10 to Grade 11 need not be as great as the change from Grade 3 to Grade 4. In fact, the change from 3 to 4 is more than twice as great as the change from Grade 10 to 11. All three scores are used on some analyses to demonstrate the different results that can be obtained by using these different scores. The nature of these scores is more fully explained in the test manual (7).

## C. RESULTS

The statistical tests employed for the analysis of the results were non-parametric in nature. Wilcoxon's (12) paired replicate test was used to test for differences between two groups of equated children and between the same children at different time periods. Wilcoxon's (12) non-parametric analysis of variance was employed to test for differences among the children over all subtests and among the subtests over all children. White's (11) extension of Wilcoxon's unpaired replicate tests for unequal $n$'s was applied to investigate differences between independent groups such as boys vs. girls, arithmetic vs. non-arithmetic pupils, schizophrenics vs. non-schizophrenics, etc. Rank order correlation coefficients were utilized to determine amount of relationship between original score and change in score, *IQ* and change, etc.

### 1. *Comparison of Intelligence of the Disturbed Children to the Norm Population*

The disturbed children under consideration here were neither superior nor inferior to normal children of their age on verbal *IQ* performance *IQ*, or full scale *IQ* ($p > .05$). Thus, results concerning the academic achievement of disturbed children cannot be ascribed to abnormally high or low intelligence. On the other hand, one must consider the fact that the median absolute difference between the verbal and performance scores is 12 with a range from 0-49.

### 2. *Statistical Tests Concerning Academic Achievement*

*a. Individual differences in academic achievement.* Significant differences ($p < .001$) were found among children for both test administrations. This fact allows us to search for factors that might explain these differences within the group (see later analyses).

*b. Differences among areas of academic achievement.* We found a statistically significant difference among the different areas ($p < .001$ for the first test, $p < .02$ for the second test). In both tests, starting with the most competent, the order of competence is word meaning, spelling, paragraph meaning, followed closely by language with arithmetic reasoning and computation tied for last place and much lower than any of the other areas. The finding that these children do worse in arithmetic computation agrees with the conclusion of Feinberg's (6) study of the atypical children referred to above. His explanation is based on his characterization of his group as emotionally disturbed "since they are children placed in an atypical situation and away from their immediate families." He concludes that disturbed children do poorly in subjects requiring exactitude in thinking. This does not stand up in this sample for spelling on which the disturbed children of this sample do next best after word meaning. The results on both test administrations shows a real break between the English and arithmetic areas. This suggests that the difference in functioning in these different areas of academic achievement might perhaps be better explained by assuming that these disturbed children have relative impairment of their abstract ability more of which is needed for arithmetic than for word meaning. It must be noted, however, that Feinberg considers that pattern of performance in which the children do better in the reading part than in the arithmetic part to be present in disturbed children. This aspect is certainly corroborated here.

c. *Comparison of the academic achievement of disturbed boys with that of disturbed girls.* No statistically significant ($p > .05$) differences were found between boys and girls in any area of academic functioning. The children were also compared on their verbal, performance and full scale intelligence with the same result, i.e., boys did not differ significantly from girls ($p > .05$). The differences in academic achievement found among the children cannot therefore be explained in terms of sex differences. At the same time, the lack of difference allows us to treat the boys and girls as one group for analyses concerning academic achievement.

d. *Comparison of academic achievement of disturbed children with that of a corresponding group of normal children.* When the second test scores of the disturbed children were compared with their corresponding normal modal-age group, they were inferior beyond the .05 level of significance for paragraph meaning and word meaning and beyond the .01 level of significance for spelling, language, arithmetic reasoning, and computation. When they were compared with their corresponding total-grade group, it was found that the disturbed children did not differ significantly ($p > .05$) from the "normal" group in paragraph meaning, word meaning, or spelling but were inferior in language, arithmetic reasoning, and computation. The same statistical test was applied in still another way, i.e., in terms of *K* scores which yield the most accurate measure of achievement of all three scores involved. This last technique shows no statistically significant ($p > .05$) difference between the disturbed children and the normal group in the area of word meaning; all other areas indicate that the disturbed children achieve less than their corresponding norm group (paragraph meaning, $p < .05$, spelling, language, arithmetic reasoning, and computation ($p < .01$). These analyses all demonstrate that the language area (really the formal aspects of language) as well as arithmetic reasoning and computation are subjects in which the disturbed children are inferior to normal children. As far as the other three subjects are concerned, the disturbed children are inferior to the modal-age group in all of them; when compared in terms of *K* scores, however, then the disturbed children do not differ significantly from the modal-age group in word-meaning ($p > .05$) although they are still inferior in paragraph meaning ($p < .05$) and spelling ($p < .01$); finally when the disturbed children are compared to the total-grade group then they do not differ from them in paragraph meaning, word meaning, and spelling.

These analyses indicate that the disturbed children are: (*a*) inferior to the modal-age group in all areas of academic functioning when the comparison is made in terms of grade scores; (*b*) inferior to the modal-age group in all areas of academic functioning except word meaning (where there is no difference) when the comparison is made in terms of *K* scores; and (*c*) inferior to the total-grade group only in language and arithmetic reasoning and computation. The difference between the grade and *K* scores reveals quite clearly that while the children may be significantly behind in their studies when measured in terms of grades, this is not so in terms of a "true" measure of achievement for the word meaning area. Comparison to the total grade group further points out that these disturbed children are inferior only when compared to the typical children of each grade, not to all children of corresponding grade; this means that at least in paragraph meaning, word meaning, and spelling, these children have not fallen behind their peers. Again the unevenness in academic functioning can be noted here, this time in terms of unevenness of retardation.

*e. Individual differences in amount of academic achievement retardation or advancement.* The children were found to differ among themselves (on basis of the second test) in amount of academic retardation or advancement ($p < .001$).

*f. Differences among academic areas in retardation and advancement.* A statistically significant difference was discovered for all three types of scores ($p < .002$) thus corroborating the finding of unevenness of functioning. The total-grade and the modal-age grade scores resulted in the same order of amount of retardation. The order in the direction of decreasing retardation beginning with the area of greatest retardation was: arithmetic reasoning, arithmetic computation, language, paragraph meaning, spelling, word meaning.

The *K* scores of each area were multiplied by the reciprocal of the range of each area (since the range of *K* scores differs from one area to another) in an effort to equate the areas. This order, which takes into consideration actual amount of material to be learned rather than grades to be passed through, goes from greatest to least amount of retardation: arithmetic computation, arithmetic reasoning, language, with word meaning and spelling tied for the same place, and last paragraph meaning. This result does not differ radically from the order arrived at by means of the other scores. The differences must, however, be taken into consideration. For both types of grade scores arithmetic computation shows just a little less retardation than arithmetic reasoning; the *K* score analysis, on the other hand, indicates that the children are by far more retarded in arithmetic computation than in arithmetic reasoning. As far as discrepancies between the two types of scores are concerned, it is suggested that the grade scores are important for actual immediate placement; for making a statement concerning the amount of learning that a given child must do to achieve a particular level of knowledge, however, the *K* scores are much superior.

### 3. *Statistical Tests Concerning Change in Academic Achievement*

*a. Significance of change in each of the six areas of academic achievement.* Modal-age grade scores indicated a statistically significant improvement in paragraph meaning, arithmetic reasoning, and arithmetic computation beyond the .01 level, in word meaning at the .02 level and in spelling and language at the .05 level. *K* scores indicated a statistically significant improvement ($p < .05$) in all areas but spelling where there is a trend ($p > .05$) in the same direction. The discrepancy between the grade score and the *K* score (which constitutes an equal interval scale) must again be interpreted to mean that while grade-wise the children may have gained more than they lost, this is not true when measured in terms of actual achievement (*K* score).

As far as the significance of the improvement in academic functioning is concerned, unfortunately we had no independent control group as such to determine whether a disturbed child who receives something called residential treatment improves more than one who does not.

*b. Comparison of the improvement in academic achievement of disturbed children with a group of corresponding (in terms of grade placement) normal children.* Using the modal-age group grade scores, we arrived at the following results: The disturbed children improved approximately as much ($p$ for difference $> .05$) as normal children correctly placed at the former's first tested grade level in spelling, language, arithmetic computation, and arithmetic reasoning. There is a non-significant trend for the disturbed children to improve less than the normals in these subjects. In paragraph meaning there is a non-significant trend in which the disturbed children improved somewhat more than the normals. In word meaning, however, the disturbed children improved significantly less ($p < .02$) than the normals during the one-year period.

Using the *K* scores, we found the same results as those cited for the modal-age group grade scores. In other words, for this particular analysis the inequality of the intervals of grade scores did not distort the substantive findings.

The fact that the disturbed children can be said not only to improve but to improve a year's worth in a period of a year may be interpreted as a sign of some general improvement in children who when coming to Linden Hill School were behind in these areas.

On the other hand, it is interesting to note that the only area in which the disturbed children did not improve as much as one would expect of normal children during a school year, namely word meaning, was also the only area where the children were functioning at about the same level as the modal-age group at the time of the second test. In other words, the disturbed children apparently did not have to improve much in order to achieve their expected grade level in word meaning in the second test, since they were originally doing well in word meaning. As for the children's progress in the other areas of academic functioning, they advanced as much as one would expect of a normally placed child. The fact that they were originally retarded, however, resulted in their remaining behind. The children's progressing at an accelerated pace is an event that has not yet occurred, but should it occur, it may certainly be regarded as a new and separate stage of progress. At this writing, the fact that these children have not continued to "not progress," but have begun again where they last left off may be viewed as improvement.

*c. Individual differences among the children in improvement (or deterioration) of academic achievement.* Analysis of the differences between the first and second test performances indicated that there are statistically significant ($p < .05$) differences among the disturbed children, i.e., all children did not improve to the same extent.

*d. Differences in improvement (or deterioration) among different areas of academic achievement.* No significant differences ($p = .70$ for modal-age scores, $p = .85$ for $K$ scores) were found among the areas of academic achievement. In other words, this analysis would seem to indicate that improvement while being least in word meaning as discussed above, is still sufficiently uniform so that differences in improvement rates from one area to another can be explained in terms of chance variation.

*c. Relationship of changes in academic achievement to the original scores.* The first hypothesis that can be tested concerning the possible reason for change is whether children who have learned little to begin with improve more than children who have already learned up to a certain point. Correlation coefficients were computed separately for each area of academic achievement in terms of grade scores and *K* scores. All correlations except for one area (with both types of scores) were not high enough to be statistically significant ($p > .05$). The area of language showed a correlation of —.55 in terms of grade scores and of —.73 in terms of *K* scores (both coefficients were significant at .01 level). These two correlation coefficients corroborate the above hypotheses, namely the lower the score the greater the change and vice versa for the area of language, or put in simpler words, the child that knew little learned much, the child who knew a lot learned little.

It is interesting to note that this relationship appears to be clearer when using *K* scores, i.e., equal intervals.

*f. Relationship of changes in academic achievement to the age adequacy of the original achievement.* Age adequacy scores were computed by taking the difference between the achieved score and the score expected of a normal child of the same age; these differences were computed in terms of *K* scores only since it has already been shown that these scores provide us with the best direct measure of actual achievement. All correlation coefficients were non-significant ($p > .05$) except the one in language which was +.61 ($p < .01$). This coefficient shows that the further the child is from where he should be, i.e., the poorer his age adequacy, the more he improves. This clarifies the correlation between original achievement and improvement shown above. It would seem to indicate that at least in the language area the improvement that took place was mostly one of "catching up" rather than of forging ahead.

*g. Relationship of changes in academic achievement to intelligence.* Change in terms of *K* scores of each area of academic functioning was correlated to the *IQ* of each child. None of the coefficients demonstrated a statistically significant relationship. This is an interesting finding, since one would expect a relationship of this kind in normal children. The fact that no such relationship is present in disturbed children is perhaps an indication of the discrepancy between the potential and actual functioning in disturbed children; it would seem to indicate that the disturbed children do not make optimal use of their intelligence.

*h. Comparison of the change in academic achievement of disturbed boys with that of disturbed girls.* In no area was there a significantly different degree ($p > .05$) of improvement or deterioration (in terms of *K* scores) between the boys and the girls. In other words, no sex difference was found in this sample of disturbed children as far as improvement in academic achievement is concerned.

*i. Relationship between diagnosis and change in academic achievement.* Because of the variety of diagnoses (as was indicated in the procedure section) which makes statistical manipulation of the data impractical, two groups were formed—one composed of those children who had been labelled "schizophrenia" or "schizophrenic reaction" and one given any other diagnoses. Children whose diagnosis has been deferred were excluded from this analysis. Before comparing the two groups on change, the following control comparisons were made: on age, *IQ*, and original (first test) academic achievement. There were no statistically significant ($p > .05$) differences between the two groups on any of these variables. It is to be noted, however, that in the intelligence comparison there is a marked tendency for the schizophrenic group to show higher intelligence.

The two groups having been shown to be approximately equal on the above variables were compared as to change. No statistically significant ($p > .05$) difference could be found for any area of academic functioning tested.

A note of caution needs to be injected here. While it is true that no difference in academic functioning could be found it must be remembered that the diagnostic grouping was of such a nature as possibly to work against such a difference. While the schizophrenic group was reasonably homogeneous, the same cannot be said of the "non-schizophrenic" group. A follow-up of the relationship that we have posited in this analysis should be conducted in a larger group of children with more children in each diagnostic group.

*j. Relationship between the course of study and change in academic achievement.* For purposes of this analysis, the children were divided into groups as indicated in the method section above. No significant differences ($p > .05$) were found in the comparisons of English vs. non-English, arithmetic vs. non-arithmetic pupils on age, *IQ*, or original performance (first test) on the achievement on the English subtests for the English comparisons or on the arithmetic subtests for the arithmetic comparisons.

Since the groups (two on English and the two on arithmetic) were approximately equal on the above variables they were then compared as to change. No statistically significant ($p > .05$) difference was found between children taking a given course of study and children not taking it. The reader is referred to the method section for a note of caution concerning this result.

## D. Discussion

Our results section shows quite clearly that at least one area of the functioning of disturbed children, namely that of academic achievement can be usefully described by means of statistical manipulation of quantitative scores. Statistical manipulation of these data allows us to recognize individual differences among the children, differences in functioning in certain areas, different relations in the several areas between normal and disturbed functioning, lack of sex difference in academic functioning and so on. One of the most interesting results, however, pertains to unevenness in the functioning of these disturbed children. This, as has already been mentioned above, is a phenomenon discovered by Feinberg (4) in his group of socially maladjusted children (delinquents). A more detailed inspection of the unevenness found in his and in this study, however, should be revealing.

We have taken out of Feinberg's study those areas of academic functioning which he tested in common with those tested in this study. We then rank ordered the tests in terms of achievement as well as in terms of retardation (difference between grade placement which was done according to age and achieved grade according to the Stanford Achievement Test). Feinberg's delinquent children showed exactly the same order of functioning in terms of achievement and retardation. Some discrepancies were found between these orders for our study. Table 1 gives us a picture of the order of academic functioning in the delinquent group and non-delinquent mentally disturbed group. Grade scores were used for ranking the children of this study to make it comparable to Feinberg's which did not employ *K* scores.

TABLE 1
Relative Order of Achievement in Different Areas as Found in Disturbed Non-Delinquent Children (This Study) and in Delinquent Children as Revealed in Feinberg's (6) Study

| Areas | Disturbed non-delinquent group | | Delinquent group | |
|---|---|---|---|---|
| | In terms of achievement | In terms of retardation | In terms of achievement | In terms of retardation |
| Word meaning | 6 | 6 | 6 | 6 |
| Spelling | 5 | 5 | 3 | 3 |
| Paragraph meaning | 4 | 4 | 4 | 4 |
| Language | 3 | 3 | 2 | 2 |
| Arithmetic reasoning | 1.5 | 1 | 5 | 5 |
| Arithmetic computation | 1.5 | 2 | 1 | 1 |

*Note:*—The higher the rank the higher the achievement or the less the retardation.

It is interesting to see that the unevenness is very much the same in the two groups with the large exception in arithmetical reasoning. Feinberg's delinquent children seem to perform least well in those areas where there is a necessity to conform to specific well-defined rules, arithmetic computation, language (mostly grammar) and spelling, but to have no problem in reasoning where alternate routes may lead to the right answer. On the other hand, the disturbed non-delinquent children of this sample show somewhat the same difficulty in doing tasks which require obeying rigid rules, but in addition manifest very poor reasoning ability. The latter is evidenced by the lowest rank being assigned to the arithmetical reasoning subtest.

The following interpretation is suggested in view of the above difference. The delinquent's poor functioning in areas requiring discipline is due to his "unwillingness" to conform, whereas his relatively good functioning in reasoning is manifest where his "unwillingness" to conform does not interfere as much with this area of functioning. We might, on the other hand, interpret the disturbed non-delinquent's poor functioning in areas requiring discipline to be due to his "inability" to function in academic areas whether it be in conforming to particular rules or whether it be in reasoning; in other words, the pathology of the disturbed non-delinquent children may be said not only to interfere with ability to follow rules but also with ability to reason, i.e., to deal rationally with the environment.

Thus more detailed inspection of unevenness in academic functioning might allow us to predict whether a given child is disturbed in a way that will manifest itself asocially (by delinquency) or by way of the mental disease symptoms. This line of thinking might well bear further investigation.

In the analysis of change the utility of quantitative treatment again becomes apparent. It is interesting to note that there was apparently no difference among the areas in the rate of improvement. While one change is not enough on which to base much of a generalization, the uniformity of change does suggest that unevenness in academic functioning may well be a constancy in these children since the change tends to preserve the unevenness of the functioning. It will be interesting to observe the character of further changes in these children to observe whether a decrease in unevenness will be correlated to a general alleviation of mental disturbances.

With one minor exception, change (whether improvement or deterioration) is found not to be related to factors such as intelligence, age adequacy of original achievement, sex, or even the course of study. This would lead us to infer that the usual relationship between change in achievement and these factors is beclouded by the effect of the pathology, and that lack of correspondence might be useful as a sign of pathology. An adequate measure or measures of pathology will eventually be set up in order to test this inference.

A general comment is fitting concerning the individual differences found in this study. The scores that differ from individual to individual may eventually serve to help us in classifying the children in meaningful ways such that more homogeneous groups result, and may eventually serve to reveal relationships which would otherwise be observed by the heterogeneity of the groups.

Finally a comment is in order on the use of *K* scores. The above analyses showed quite clearly that one can obtain different results depending upon the type of score used. The fact that *K* scores appear to be the best indicators of actual achievement leads to the conclusion that in future studies of academic achievement use should be made of this type of score.

## E. SUMMARY

In this study an attempt was made to examine some of the facets of academic achievement in mentally disturbed adolescents. The basic data consisted of the Stanford Achievement Test scores of 26 such children who were tested twice with a one-year interval between testing sessions.

The following results were derived from the data (these were largely based on *K* scores):

1. No significant difference was found between the disturbed children and normal children in verbal, performance, or full scale *IQ*.
2. Significant individual differences among the disturbed children were found in academic achievement (for both test sessions), in age adequacy of academic achievement (test applied only to the second session), and in change in academic achievement.
3. Differences among areas of academic achievement were found to be statistically significant (from highest to lowest-word meaning, spelling, paragraph meaning, language, with arithmetic reasoning and computation tied for last place), and for age adequacy scores (from most to least adequate-paragraph meaning, with word meaning and spelling tied for next place, language, arithmetic reasoning and arithmetic computation), but not for change, i.e., change was the same in all areas within random errors.
4. Comparison of the disturbed children to a corresponding group of normal children showed that the disturbed children achieved significantly less in all areas of achievement except in word meaning.
5. A statistically significant improvement was found in all areas of

academic functioning except in spelling where a trend in the same direction was evident.

6. Comparison of the disturbed children to a corresponding group of normal children indicated that the disturbed children improved approximately as much as normal children in spelling, language, arithmetic computation, and arithmetic reasoning (although there was a nonsignificant trend for the disturbed children to improve less than the normal children) and in paragraph meaning (where the nonsignificant trend showed that the disturbed children improved somewhat more than the normal children); in word meaning, however, the disturbed children improved significantly less than the normal children.

7. An attempt was made to examine the relationship between change in academic achievement and the following six variables: original score, age adequacy of original score, intelligence, sex, diagnosis, course of study.

These variables were shown to be largely unrelated to change; the only exceptions were the relationship between original score in language and change in language (rho $= -.73$, $p < .01$) which indicated that the less the child knew originally the more he improved and the relationship between age adequacy in language and change in language (rho $= +.61$, $p < .01$) which indicated that the greater the retardation in a given child, the greater was the improvement.

The following inferences were made on the basis of the results:

1. Delinquent children were hypothesized to differ from non-delinquent disturbed children in that the former do not conform to rigid rules while the latter have difficulties with reasoning ability.

2. Since the disturbed children manifest a consistent pattern of unevenness among the areas of academic functioning and since the change in achievement is even over all the areas so that the unevenness of performance is preserved, we might hypothesize the particular pattern of functioning as a constancy in particular types of children (disturbed non-delinquent ones as against delinquent ones). Furthermore, we might suggest amount of unevenness as a possible index of amount of disturbance.

3. The lack of relationship between change in academic achievement and sex, intelligence, age adequacy of original achievement, and even the course of study leads us to suggest amount of such relationship in any one child as another possible index of disturbance.

4. The different results obtained by using different types of scores suggest the importance of using *K* scores (whenever differences in achievement

are under consideration, whether within individuals or between them) as the most accurate measure of actual achievement.

## REFERENCES

1. BETTELHEIM, B. Love Is Not Enough: The Treatment of Emotionally Disturbed Children. Glencoe, Ill.: Free Press, 1950.
2. ———. Truants From Life. Glencoe, Ill.: Free Press, 1955.
3. DELANY, L., HOTKINS, A. S., KORNBERG, L., SALZINGER, K., & SCHERER, M. Design of a treatment setting for atypical adolescents. Paper delivered at the annual meeting of the Jewish Board of Guardians, 1955.
4. FEINBERG, H. Achievement of a group of socially maladjusted boys as revealed by the Stanford Achievement Test. *J. Soc. Psychol.*, 1947, **26**, 203-242.
5. ———. Achievement of a group of children in foster homes as revealed by the Stanford Achievement Test. *J. Genet. Psychol.*, 1949, **75**, 293-303.
6. ———. Achievement of children in orphan homes as revealed by the Stanford Achievement Test. *J. Genet. Psychol.*, 1954, **85**, 217-229.
7. KELLEY, T. L., MADDEN, R., GARDNER, E. F., TERMAN, L. M., & RUCH, G. M. Stanford Achievement Test. New York: World Book, 1953.
8. KORNBERG, L. A Class for Disturbed Children: A Case Study and Its Meaning for Education. New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1955.
9. REDL, F., & WINEMAN, D. Children Who Hate: The Disorganization and Breakdown of Behavior Controls. Glencoe, Ill.: Free Press, 1951.
10. ———. Controls From Within: Techniques for the Treatment of the Aggressive Child. Glencoe, Ill.: Free Press, 1952.
11. WHITE, C. The use of ranks in a test of significance for comparing two treatments. *Biometrics*, 1952, **8**, 33-41.
12. WILCOXON, F. Some Rapid Approximate Statistical Procedures. New York: American Cyanamid, 1949.

*Biometrics Research*
*New York State Psychiatric Institute*
*722 West 168th Street*
*New York City*

*The Journal of Genetic Psychology*, 1957, **90**, 255-258.

# PROGRESS IN THE USE OF ENGLISH AFTER 22 YEARS BY CHILDREN OF CHINESE ANCESTRY IN HONOLULU

*Department of Psychology, University of Hawaii*

---

MADORAH E. SMITH

---

The present study[1] is a repetition of a portion of a study of bilingual children in Hawaii (4) completed nearly 20 years ago. At that time the children in Hawaii were found to be much retarded in their language behavior. Since it is now 22 years since the last sample of children of Chinese ancestry was recorded, it seemed desirable to repeat that part of the study in order to see what had happened during that time. This time only four- and five-year-old children were used. The present sample of children was selected so that their fathers' occupations would agree in percentage distribution with those of men of the same ancestry in Honolulu according to the 1950 census. The 1930 census data had been so used before and it was found that there has been considerable upward social mobility for the percentages in higher levels was greater and the Barr ratings of fathers' occupations has risen from 9.62 to 10.46. That the children of 1955 hear much more English is clear from the language ratings of their homes made on a five-point scale according to amount of English used, for the ratings have risen from 2.5 or about an equal amount of both languages heard to 4.2; very little Chinese being used unless a grandparent was living with the family and even then some grandparents preferred to use English. In 1933, only 75 per cent of the children's parents had been born in the United States; in 1955, all of them had. They had also completed on the average twice as many years of schooling and no longer called any of the children by a Chinese given name, whereas 76 per cent of our subjects in 1933 had been so called.

Even though as much Chinese as English was used in their homes in 1933, most of the children preferred to use English when at play (the condition under which the records were taken) and on the average used 82 per cent English words but now that average is 98.7 per cent and only one child used more than eight Chinese words. Of the Chinese words used by the other 49 children almost one-fourth were proper nouns, one-fifth

---

*Received in the Editorial Office on February 10, 1956.

[1] This study was made possible through a research grant by the Hawaii Psychological Association and the aid of Miss Florence Chang.

were words referring to grandparents and the only other word used by as many as three children was "koko" or elder brother, a word for which English has no exact equivalent. In 1933 most of the children confused the two languages, using both unnecessarily in the same sentence as was the case in 17 per cent of their sentences, but in 1955 they used only 2.4 per cent of such sentences or 2.1 per cent if the sentences in which "koko" occurred are not included.

TABLE 1
COMPARISON OF BACKGROUND AND LANGUAGE BEHAVIOR OF HONOLULU CHILDREN OF CHINESE ANCESTRY IN 1933 AND 1955

| | 1933 | | 1955 | | | Difference | |
|---|---|---|---|---|---|---|---|
| Number of children | 51 | | 50 | | | | |
| % of parents born in U. S. | 75 | | 100 | | | | 25 |
| Education of midparent in years | 6.1 | | 12.3 | | | | 6.2 |
| Barr Rating of fathers' occupations | 9.62 | | 10.46 | | | | 0.84 |
| Language rating of homes | 2.5 | | 4.2 | | | | 1.7 |
| % of children called by English names | 76 | | 100 | | | | 24 |
| % of words spoken in English | 82 | | 98.7 | | | | 16.7 |
| % of sentences that were mixed | 17 | | 2.4 | | | | —14.6 |
| % of complex and compound sentences | 9 | | 11 | | | | 2 |
| % of verbs conjugated | 16 | | 28 | | | | 12 |
| At age | 4 | 5 | 4 | 5 | | 4 | 5 |
| Mean age in months | 47.2 | 60.5 | 47.6 | 59.6 | | 0.4 | —0.9 |
| Sentence length | 3.4 | 3.8 | 4.03 | 4.65 | | 0.63 | 0.85 |
| σ of av. | | | | | σ of Diff. | | |
| Mean age in months | 47.2 | 60.5 | 47.6 | 59.6 | | 0.4 | 0.9 |
| Sentence length | 3.4 | 3.8 | 4.03 | 4.65 | | 0.63 | 0.85 |
| σ of mean | .07 | .13 | .17 | .15 | σ of Diff. | 0.18 | .20 |
| *t* | | | | | | 3.5* | 4.5* |
| Error Index | 332 | 320 | 218 | 171 | | 114 | 149 |
| σ of mean | 20 | 18 | 17 | 13 | σ of Diff. | 26 | 22 |
| *t* | | | | | | 4.4* | 5.8* |
| Different words per 100 | 53.60 | 58.15 | 57.84 | 59.64 | | 4.24 | 1.49 |
| σ of mean | 1.52 | 1.11 | 1.19 | 1.26 | σ of Diff. | 2.2 | 1.7 |
| *t* | | | | | | 1.9 | 0.87 |

*Significant beyond the 1 per cent level of confidence.

In 1933 the children used very immature language with short sentences, most of which were simple in form with very little inflection of words and very many errors. Even though words in both languages were counted when they might be identical in meaning, the *DW*, or number of different words per 100 running words, was found to be somewhat lower than the 55.8 and 59.4 calculated for samples of 50 each of monolingual children of four and five years respectively. The sentences which in 1933 were only slightly longer than the three-year-old average for monolingual children (3) in 1955 had nearly reached the averages of 4.3 and 4.7 for their ages. In *DW* they slightly but not significantly excel the monoglots at both ages. They now used only one per cent fewer complex and compound sentences than the average monoglot at four years (1); however, pidgin English has not dropped out of use as much as Chinese has. So in errors peculiar to pidgin there has not been so much improvement. Only the five-year-olds attain the monolingual norm for three-year-olds as to per cent of verbs conjugated either by change of form or by the use of the correct auxiliary. Nor at neither age do they make as few errors as the average three-year-old monoglot. Nevertheless in the 22 years, there is a gain of 75 per cent in the number of verbs conjugated and the error index (2) has been almost halved by five years.

Only for the more important measures were the $\sigma_{av}$'s calculated. The gains made in the 22 years for both four- and five-year-olds are highly significant for both Error Index and Sentence Length. Only for the four-year-olds does the gain in *DW* approach significance.

Although pidgin English caused the greatest amount of error, 20 per cent of our 1955 sample used not a single pidgin phrase. Of the 40 such errors used in 1933, only 26 were still being used in 1955 and only four of them by more than one-fifth of the children. These four were "no" for "not," "one" for "a," "go" as an auxiliary verb, and "get" for "have" or "there is."

## Summary and Conclusions

1. Fifty children of Chinese ancestry in Honolulu who were from 36 to 60 months in age are compared as to their use of the English language with the same number of children of like age and ancestry who had been studied 22 years before. The children at both periods were selected so that the distribution of their fathers' occupations should agree with that given in the latest census.

2. In 1933, the children had been found to be much retarded in their use of English even though they preferred to use it rather than the Chinese which was spoken half the time in their homes. So 82 per cent of the words they used were English.

3. In 1933, 75 per cent of the parents had been born in the United States, 24 per cent of the children were called by Chinese names, the Barr Ratings of fathers' occupations averaged 9.62, language ratings of homes according to amount of English used averaged 2.5 and the education of

mid-parent 6.1 years. In 1955, all the parents had been born in the United States, all children were called by English names, Barr Ratings had risen to 10.46, language ratings to 4.2, and education of mid-parent to 12.3 years. So the children had a better background for learning English.

4. In 1955, only one child could be considered bilingual and the children used on the average 98.7 per cent English words. Mixed sentences had fallen from a frequency of 17 per cent to 2.4 per cent. The only Chinese words used by more than three children were those by which they called their grandparents.

5. Save for *DW*, which was not much below monolingual norms, on all measures of language proficiency, the children in 1933 barely attained or failed to reach the norm for three-year-old monoglots. The children in 1955 had almost attained the norms for their ages in sentence length, that of the four-year-old average in use of complex and compound sentences, and in *DW* only a little below norm at four years and very slightly above at five years. The gain in sentence length is a highly significant one at both ages.

6. Although the gains in number of verbs conjugated and in reduction of errors are very great, they are still not enough to more than just attain the norm for three-year-old monoglots as to conjugation of verbs and not to reach it in the case of errors. This is mainly due to the contamination of the English they hear by pidgin.

7. To conclude, the preschool children of Chinese ancestry in Honolulu are now with few exceptions no longer bilingual and are exposed to very little Chinese; but their speech is still contaminated by pidgin English so that they still make too high a per cent of errors for their age. Although there has been marked gain in the last 22 years so that the task of teaching correct English will be less difficult than in 1933 for teachers in Hawaii, it will still be a problem.

## REFERENCES

1. SMITH, M. E. An investigation of the development of the sentence and the extent of vocabulary in young children. *Univ. Iowa Stud.: Stud. Child Welf.*, 1926, **3**, No. 5. Pp. 92.
2. ———. Grammatical errors in the speech of preschool children. *Child Devel.*, 1933, **4**, 183-190.
3. ———. A study of some factors influencing the development of the sentence in preschool children. *J. Genet. Psychol.*, 1935, **46**, 182-212.
4. ———. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among preschool children of non-American ancestry in Hawaii. *Genet. Psychol. Monog.*, 1939, **21**, 119-284.

*1533 C Wilder Avenue*
*Honolulu 14, Hawaii*

*The Journal of Genetic Psychology*, 1957, **90**, 259-266.

# LATENT EXTINCTION IN A MULTIPLE-*T* MAZE WITHIN HETEROGENEOUS AND HOMOGENEOUS ENVIRONMENTS*

*Department of Psychology, University of Texas*

W. LYNN BROWN[1] AND EDWARD S. HALAS

## A. PROBLEM

Recent investigations from this laboratory by Brown and Halas (1, 2) on terminal extinction in heterogeneous and homogeneous environments have shown a deterioration in performance following extinction in both time and error measures. Under limited response extinction, i.e., extinction only on the terminal unit and in the goal-box, some *S*s were extinguished on a white terminal unit and goal-box and others on a black unit and goal-box in a heterogeneous environment (1), while others were extinguished in a homogeneous environment under the same maze conditions (2). Following eight extinction trials under each condition, *S*s were then run on the original maze and were checked for both time and errors. In the heterogeneous environment the *S*s which were extinguished on the training terminal unit and goal-box (white) made more errors on a test run following extinction than *S*s which were extinguished on a black unit and goal-box which was new to the *S*s. Likewise the *S*s took more time on the test run. Brown and Halas discussed the results in terms of secondary reinforcement and concluded that "in the present experiment a reduction in secondary reinforcing cues strengthened resistance to extinction."

In the investigation in which a homogeneous extra-maze environment was maintained, Brown and Halas (2) found the same pattern of deterioration in performance but with fewer errors and faster running time following extinction. However, the performance on the test run following extinction differed from that of the last run before extinction beyond the 1 per cent level. Brown and Halas concluded "that the difference in performance between the *S*s in the two experiments is attributable to differences in extra-maze stimuli."

Since the *S*s responded differently following extinction training on a test trial in heterogeneous and homogeneous environments, one might assume that a further reduction in stimuli and responses associated with extinction would

---

*Received in the Editorial Office on March 12, 1956.

[1] This investigation was supported in part by a grant to the senior author by the Research Institute of the University of Texas.

reduce still more the deterioration in both time and errors on the test trial. In the present experiment on latent extinction the *S*s were placed in the goal-boxes only during the extinction period, and they were not permitted to run the final unit.

The hypothesis of "latent extinction" finds confirmation in the experimental results of Seward and Levy (8) and Deese (5). Negative findings have been reported by Bugelski, Coyer, and Rogers (3) and also by Scharlock (7). More recently Moltz (6) found that latent extinction is "dependent on the reduction or elimination of the secondary reward value of cues present in the goal-box during the latent extinction period" (p. 399). Moltz also reduced extra-maze cues by completely surrounding the experimental apparatus with several layers of cheesecloth. This design tended "to provide a nearly uniform perceptual field."

In an investigation in this laboratory Crawford (4) found, while using an elevated *T*-maze, that "latent extinction effects are to be expected when the nature of the pre-extinction experience is such that it represents a discriminable change from the conditions of training." Crawford interpreted his findings "in terms of the reduction of secondary reward value as a consequence of nonreinforced components of the goal response and in terms of sign-learning."

## B. MATERIALS AND METHODS

### 1. *Experimental Rooms*

The present experiment on latent extinction was conducted in the two experimental rooms used by the former studies on "limited response extinction" (1, 2). The room with a heterogeneous environment is 13 ft. 4 in. x 16 ft. 9 in., and has a 9-ft. ceiling. One window is centered on the west wall and one on the east. Two doors are on the south and one on the north. The room, designed to give a homogeneous perceptual field, has an experimental dome 18 ft. in diameter with 5½-ft. side walls arched to an 8-ft. center. The dome is covered with a fine mesh screen painted silver gray, thus giving a homogeneous surface radially symmetrical about the geometric center. Indirect lighting was used during the experiment.

### 2. *Subjects*

Thirty-six experimentally naïve white female rats, ranging in age from 120 to 150 days old, were used. The *S*s were divided into two groups of 18 each, with Group *C* receiving training and latent extinction in the heterogeneous environment and Group *D* in the experimental dome.

### 3. *Apparatus*

The 14-unit multiple-$T$ maze used in the two former studies was used in the present experiment. The first 13 units were painted mid-gray while the 14th and goal-box were painted white.

### 4. *Maze Training*

The training procedure was the same as in the former studies. *S*s were given one run a day for nine days. When an *S* entered the goal-box, the terminal unit was moved out of position to prevent retracing. Each *S* was left in the goal-box upon the completion of the run for one minute and was allowed to eat of a wet mash, after which the *S* was returned to its home cage and fed one large pellet of Purina Laboratory Chow.

### 5. *Extinction Procedure*

On the tenth day, the day after maze training, the *S*s in both Groups *C* and *D* were divided into two matched subgroups of nine *S*s each on the basis of time and error score. The *S*s in each of the four subgroups were given eight extinction trials. For the *S*s in Group *C*, the *S*s trained in a heterogeneous environment, nine (Subgroup *C-w*) were given extinction periods in a white goal-box (the box used during training) and nine (Subgroup *C-b*) were given extinction periods in a black goal-box. An extinction period consisted in placing the *S* in the empty goal-box for one minute, after which the *S* was picked up and held for 15 seconds. After being held for 15 seconds, the *S* was placed back in the empty goal-box for the second extinction period. At the conclusion of eight extinction periods, the *S* was held for one minute before being placed at the beginning of the training maze for the test run.

The *S*s in Group *D*, the group trained in a homogeneous environment, were likewise divided into two matched subgroups of nine each. The *S*s in Subgroup *D-w* were given extinction in a white goal-box and those in Subgroup *D-b* in a black goal-box. The extinction procedure was the same as for Group *C*.

## C. Results and Discussion

The performance curves for time and error are presented in Figures 1 and 2 for both groups. An examination of the results as presented will show that the groups performed in a similar manner on both measures and that they terminated on the ninth day, the day before extinction, at about the same level. On the ninth day of training Subgroup *C-w* had a running time average of 19.44 seconds and an average of 51.44 seconds on the test run after eight extinction trials on the 10th day. This difference in performance is significant beyond the 1 per cent level ($t = 5.94$). Subgroup *C-b* had a running time average on the ninth day of 19.66 seconds and an average of 25.44 after extinction. This difference is significant also at the 1 per cent level ($t = 3.33$). The difference between the two subgroups was likewise significant beyond the 1 per cent level ($t = 4.44$).

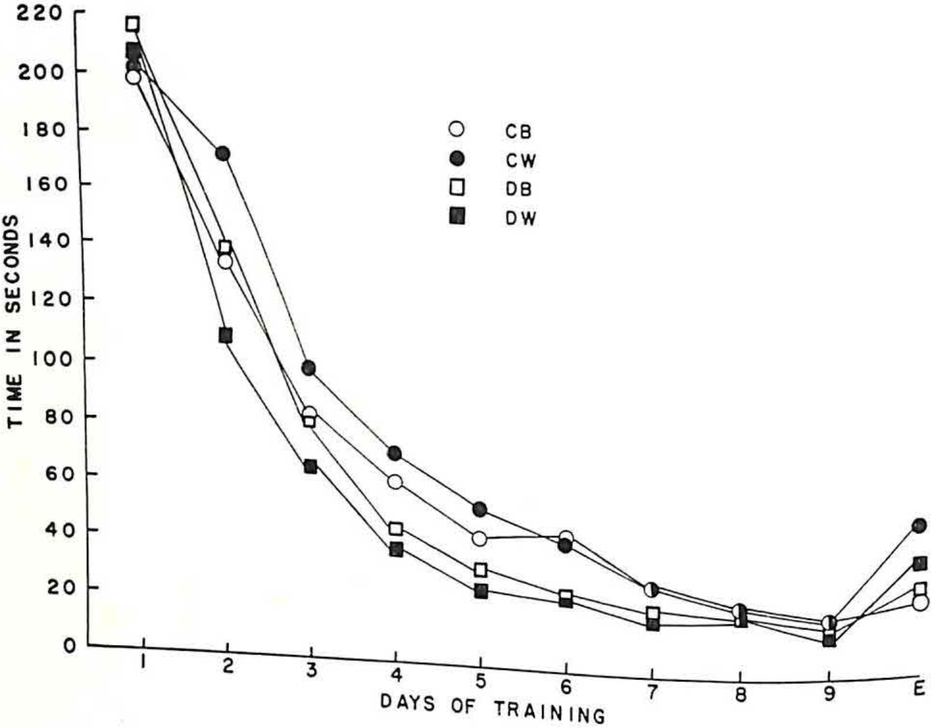

FIGURE 1
Running Time During Learning and Extinction

Error differences between the subgroups of Group *C* were in the same direction as were those for time. Subgroup *C-w* made one error on the ninth day of training and 23 on the test run following extinction. This difference is significant beyond the 1 per cent level ($t = 3.69$). Subgroup *C-b* also made one error on the ninth day of training but only seven on the test run. A $t$ of 1.81 here indicates that the difference is not significant. Finally, when the two subgroups are compared on the test trial, they just miss being significant at the 5 per cent level ($t = 2.24$).

The *S*s in Group *D* were trained in a homogeneous dome and were also given extinction training and a test trial in the same room. Subgroup *D-b* had a running time average on the ninth day before extinction on the 10th of 16.88 seconds and 26.77 after extinction. This difference is significant beyond the 1 per cent level ($t = 5.74$). Subgroup *D-w* had a running time average of 16.22 on the last day of training and 38.22 following extinction. This difference is also significant beyond the 1 per cent level ($t = 3.57$). The time difference between the groups on the test trial was not significant ($t = 1.63$).

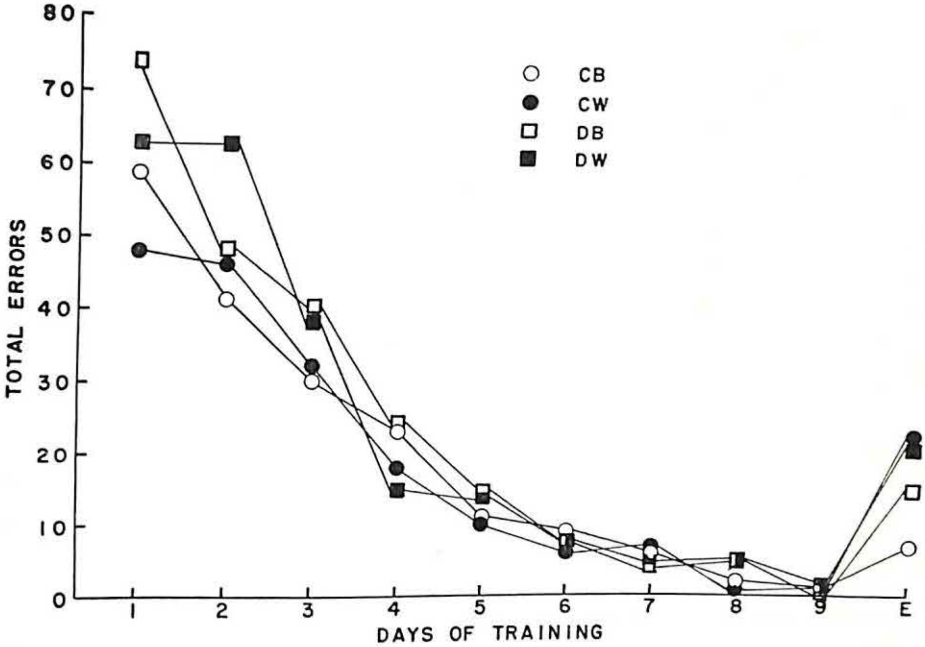

FIGURE 2
Curves for Learning and Extinction

When the subgroups are compared for errors following extinction, we find the difference in the same direction as was found for time. Subgroup *D-b* made one error on the ninth day of training and 15 on the test trial following extinction, a difference significant beyond the 5 per cent level ($t = 2.46$). Subgroup *D-w* made no error on the final day of training and 21 following extinction. This difference is very significant beyond the

1 per cent level ($t$ =5.10). The difference in errors between the subgroups on the test trial was not significant ($t$ = .73).

The same pattern of performance was found in the results obtained by Subgroups *D-w* and *D-b* of Group *D* as was observed in Subgroups *C-w* and *C-b* of Group *C*. In fact, the subgroups which were extinguished on black, whether in a heterogeneous or homogeneous environment of the "limited response" type of extinction (1, 2) or following latent extinction (present study), always made fewer errors and ran faster on test trials than did the subgroups extinguished on white.

A number of comparisons may be made between the two former studies of limited response extinction and the present study on latent extinction in terms of error scores. (*a*) The *S*s in the studies using the limited response type of extinction consistently made more errors than did the *S*s on the same conditions but with latent extinction. The *S*s using limited response extinction made a total of 126 errors and those using latent extinction made 66. (*b*) The error score of the *S*s may be considered with reference to extinction on white and on black. In the first study the *S*s extinguished on white made 43 errors following extinction and those of the second made 36. Both studies involved limited response extinction, the first being in a heterogeneous surrounding and the second in a homogeneous one. When the same environmental conditions are considered for the *S*s in the latent extinction study, we find that the *S*s in the heterogeneous environment made 23 errors and those in the homogeneous made 21. When we compare the errors of *S*s extinguished on white in heterogeneous surroundings, we find that the *S*s using limited response extinction made 43 errors and those using latent extinction made 23. Likewise if we compare for homogeneous surroundings, we get error scores of 36 and 21. The comparisons above show that *S*s extinguished on white make more errors following extinction in a heterogeneous than in a homogeneous environment and likewise more errors following limited response extinction than latent extinction.

The pattern of errors for the *S*s extinguished on black differs from these extinguished on white at only one point. Again the *S*s using limited response extinction made more errors than *S*s on the same conditions but with latent extinction. The former *S*s made a total of 47 errors and the latter 22. When we compare the error scores for extinction on black, while using a limited response type of extinction procedure, we find that the *S*s made 29 errors in heterogeneous and 18 in homogeneous surroundings. When the same surroundings are considered for the *S*s using latent extinction procedures, we find that the *S*s made 7 errors in heterogeneous and 15 in homogeneous surroundings. This is the only reversal in pattern. When we compare the errors of *S*s extinguished on black in heterogeneous surroundings, we find that those using limited response extinction made 29 errors and those using latent extinction made 7. Likewise if we compare for homogeneous surroundings, we get error scores of 18 and 15. The above comparative analysis of *S*s extinguished on black shows that more errors follow limited response extinction on test runs than on latent extinction, whether the surroundings are heterogeneous or homogeneous. The *S*s extinguished on white gave the same result. Since the only difference between the studies on limited response extinction and the present one on latent extinction was in the nature of extinction procedure, it therefore appears that when stimuli and responses are reduced in the reward area during extinction, the extinction is not as complete as when the responses take place.

A comparison of all subgroups on black and white for all experiments will show that the *S*s extinguished on black made fewer errors following extinction than those extinguished on white. When the training condition is changed from white to black for extinction, the *S*s show less effect of extinction. Apparently, when using the experimental design and conditions of the present series of experiment, a "reduction of secondary reinforcing cues strengthened resistance to extinction."

An analysis of average running times before and after extinction for the subgroups of the three experiments under different surroundings and extinction procedure reveals a pattern like the one presented above for error. All subgroups in the three experiments had a longer running time following extinction than before extinction. Running time averages following extinction may be compared in several different ways. (*a*) The subgroups extinguished on white always ran slower than subgroups extinguished on black under the same extinction procedure and environmental surroundings. (*b*) The subgroups subjected to limited response extinction always ran slower than their corresponding subgroups subjected to latent extinction.

## D. Summary and Conclusions

Two groups of experimentally naïve white rats were trained on a multiple-*T* maze with the first 13 units painted a mid-gray and the final unit and goal-box painted white. Group *C* (*N* — 18) was trained and extinguished in a room with heterogeneous features while Group *D* (*N* — 18) was trained and extinguished in a homogeneous dome which gives a "uniform perceptual field." Both groups were divided into matched subgroups of nine each before extinction trials. The *S*s in Group *C* were divided in

Subgroups *C-w* and *C-b*, while Group *D* was divided in Subgroups *D-w* and *D-b*. The *S*s in Subgroup *C-w* were given eight extinction periods of one minute each in the white goal-box used during training on the 10th day of the experiment, after which they were given a test run on the original maze. The training and etxinction procedure was the same for Subgroup *C-b* as was used for Subgroup *C-w* with the exception of a black goal-box which was introduced for the extinction periods. The same pattern of extinction was used for Subgroups *D-w* and *D-b* in the homogeneous dome room. The *S*s extinguished on white (Subgroups *C-w* and *D-w*) made more errors and had a longer average running time than the *S*s extinguished on black.

The results of the present experiment on latent extinction are discussed in relation to the earlier experiments on limited response extinction (1, 2). The *S*s made fewer errors and had shorter running times following latent extinction than they did following the type of limited response extinction used in the earlier studies.

## REFERENCES

1. BROWN, W. L., & HALAS, E. S. Terminal extinction in a multiple-*T* maze within a heterogeneous environment. *J. Genet. Psychol.*, 1957, **90**, 89-95.
2. ———. Terminal extinction in a multiple-*T* maze within a homogeneous environment. *J. Genet. Psychol.*, 1957, **90**, 97-101.
3. BUGELSKI, B. R., COYER, R. A., & ROGERS, W. A. A criticism of pre-acquisition and pre-extinction of expectancies. *J. Exper. Psychol.*, 1952, **44**, 27-30.
4. CRAWFORD, F. T. Latent extinction as a function of pre-extinction experience. Ph.D. dissertation, Univ. Texas, 1956.
5. DEESE, J. The extinction of a discrimination without performance of a choice response. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 362-366.
6. MOLTZ, H. Latent extinction and the reduction of secondary reward value. *J. Exper. Psychol.*, 1955, **49**, 395-400.
7. SCHARLOCK, D. P. The effects of a pre-extinction procedure on the extinction of place and response performance in a *T*-maze. *J. Exper. Psychol.*, 1954, **48**, 31-36.
8. SEWARD, J. P., & LEVY, N. Sign learning as a factor in extinction. *J. Exper. Psychol.*, 1949, **39**, 660-668.

*Department of Psychology*
*University of Texas*
*Austin 12, Texas*

# SHORT ARTICLES AND NOTES

*The Journal of Genetic Psychology*, 1957, **90**, 267-270.

# DRIVE IN ILLNESS: AN EXPERIMENTAL STUDY OF BARRIER BEHAVIOR OF (ALLOXAN) DIABETIC AND NONDIABETIC ANIMALS*[1]

*School of Medicine, University of Oklahoma*

WILLIAM SEEMAN AND ROBERT C. GREEN, JR.

## A. PROBLEM

Although experimentally induced diabetes has been studied in animals before, the investigators who have conducted the studies have not been interested in the drive aspects of the animal behavior. The present study was undertaken to determine experimentally whether diabetic animals (in which the diabetes was induced by intravenous injections of alloxan) under specified conditions of water deprivation would differ from control animals in crossing an electric grill.

## B. APPARATUS

This apparatus was an adaptation of that described by Miller. It consisted essentially of a long, rectangular box divided into several chambers. An animal was placed in Chamber *a*, by means of a door opening at the top. The floor of Chamber *b* was an electrified grill made of 18-gauge copper wire, the strands of which were ¼ inch (0.64 cm.) apart. Resistance of 500,000 ohms was placed in the circuit, and the voltage was 500. The animal had to pass through Chamber *c* to reach the water in the goal-box, *d*. The floor, the ceiling, and one side of the box consisted of ¼ inch plywood. The fourth side was transparent plastic, so that the animal could be seen in each chamber.

## C. PROCEDURE

Eleven albino rats, approximately three months of age, made up the experimental group, and an equal number of animals were used as controls. Each animal, when not in the apparatus, was kept in a metabolic cage which permitted collection of the daily urine so that it might be tested for its sugar content. The method of inducing alloxan diabetes which we em-

---

*Received in the Editorial Office on November 23, 1955.

[1] This experiment was conducted while both authors were at the Mayo Clinic.

ployed first was a modification of that suggested by Lazarow and Palay. Subsequently however it was found more feasible to anesthetize the animals and to inject the alloxan via the saphenous vein. Thus, the method of inducing diabetes finally settled upon was intravenous injection, with a No. 26 hypodermic needle, of an aqueous solution of 2 per cent alloxan monohydrate (40 mg. per kilogram of body weight). For purposes of rigorous control the nondiabetic animals also were anesthetized although they were not, of course, injected at all. A concentration such as has been mentioned produced diabetes in all of Lazarow and Palay's animals and in all of our experimental animals as well. All experimental animals tested 24 hours after injection were found to have large concentrations of sugar (2 per cent or greater) in the urine, in addition to acetone.[1] The control animals, when tested similarly, gave no evidence of glycosuria or acetonuria as was to be expected. Both intake of water and output of liquid by the injected animals markedly increased within 24 hours of the time of injection.

All animals of both the experimental and the control group received similar treatment and training. They were first allowed a period of adaptation, in which they could become accustomed to handling and could explore the apparatus. They also were placed on a water rhythm; that is, they were deprived of water for 24 hours; then they were placed in Chamber *a* of the apparatus, in the goal box of which water had been placed. All learned to run from Chamber *a* to the goal-box. Food in plenty was always available in their home cages. During this training period the animals were run once a day for seven days, and all ran quickly to the goal-box to drink. After animals had received alloxan, water was made available to them for 24 hours. Then they were deprived of water for 24 hours. Next, they were run in the apparatus.

On the ninth day shock was introduced for the first time and the animals were allowed five minutes for crossings. If an animal crossed the grill he was permitted to attain the goal box but scarcely to drink, and then was placed again in Chamber *a*, to run again if he chose. Both crossings and "attempts" were tallied; an attempt was recorded whenever an animal went far enough to receive a shock and then recoiled. All animals were first run after being deprived of water for 24 hours. The control animals then were run again after 48 hours of deprivation of water, but the experimental animals could not be run under this condition since many of them had died.

---

[1] This condition of severe glycosuria and acetonuria persisted in all experimental subjects throughout the duration of the experiment.

## D. RESULTS

In Table 1 the numbers of crossings and the number of trials are presented for both control and the experimental animals. The results are comparable. The mean number of crossings for the experimental animals was 1.45 and for the controls 1.60. This is not a statistically significant difference. The mean number of attempts for the experimental animals was 3.63 and for the controls, 3.0, a negligible difference. The median test also was applied and it, too, yielded results which statistically were not significant.

TABLE 1

NUMBER OF CROSSINGS AND NUMBER OF ATTEMPTS TO CROSS MADE BY (ALLOXAN) DIABETIC ANIMALS AND BY CONTROL ANIMALS

| Animal number | Experimental | | Animal number | Control | |
|---|---|---|---|---|---|
| | Cross | Attempts | | Cross | Attempts |
| 1 | | 5 | 1 | 14 | 2 |
| 2 | | 5 | 2 | | 4 |
| 3 | 1 | 6 | 3 | | 3 |
| 4* | | 0 | 4 | 1 | 3 |
| 5 | 1 | 3 | 5 | 1 | 1 |
| 6 | | 3 | 6 | | 4 |
| 7 | | 5 | 7 | | 2 |
| 8 | | 4 | 8 | | 3 |
| 9 | 1 | 5 | 9 | | 6 |
| 10 | 11 | 1 | 10 | | 2 |
| 11 | 2 | 3 | 11† | | |

*Animal appeared quite ill, extremely lethargic, and merely sat in the apparatus.
†Animal died under anesthesia.

Five animals in the experimental group and threee in the control group made crossings; thus, for both groups the shock operated as an effective deterrent.

What factors account for the failure to demonstrate a significant difference in behavior between the two groups we cannot say. Certainly, as has been noted, the experimental animals, when permitted to take water, markedly increased their intake of it and their output of urine increased. However, after injection their manner of moving was a good deal more lethargic than that of the control animals and, indeed, more lethargic than was characteristic of themselves prior to the injection of alloxan. Whether this lethargy was a compensating factor, to offset the increased loss of water, we cannot say, of course.

## E. SUMMARY

Two groups of animals were trained to find water after they had run down a long, rectangular box. One group was then diabetic by means of alloxan injection. Members of both groups were then required, in their path to water, to cross an electrically charged grid, the shock from which acted as a barrier. As between diabetic and control animals, there were no significant differences in the number which crossed the barrier, in the mean number of crossings, or in the number of attempts to cross.

## REFERENCES

1. Bailey, C. C., Bailey, O. T., & Leech, R. Alloxan Diabetes. *Proc. Amer. Diabet. Assoc.*, 1946, **6**, 373-387.
2. Gomori, G., & Goldner, M. G. Production of diabetes mellitus in rats with alloxan. *Proc. Soc. Exper. Biol. & Med.*, 1943, **54**, 287-290.
3. Lazarow, A., & Palay, S. L. The porduction and course of alloxan diabetes in the rat. *J. Lab. & Clin. Med.*, 1946, **31**, 1004-1015.
4. Miller, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. Exper. Psychol.*, 1948, **38**, 89-101.

*School of Medicine*
*University of Oklahoma*
*800 Northeast 13th Street*
*Oklahoma City 4, Oklahoma*

*The Journal of Genetic Psychology*, 1957, **90**, 271-274.

# GRADUATE SCHOOLS OUGHT TO TRAIN TEACHERS OF PSYCHOLOGY*

*Department of Psychology, Hunter College*

VIRGINIA M. STAUDT

In recent years, prompted by a variety of reasons including licensing and certification legislation, many psychologists have been reflecting upon effective means of creating better understanding of psychology among the public. One obvious area for instituting improvement has been overlooked frequently in many quarters, namely, the improvement of undergraduate instruction in psychology. Such improvement must begin at the beginning, that is, with the improvement in the preparation of teachers of undergraduate psychology.

Each year thousands of students take the one-and-only course in psychology which they may ever take, usually the required introductory or general psychology. In this course a powerful influence for good or for bad may be exerted. It is, therefore, of prime importance that college teachers of courses such as this be well trained. College students should be able to find in their instructor of psychology not simply a research specialist, someone who has earned the M.A. or Ph.D. degree, or someone who is merely teaching to fill in time until the degree is obtained or until a more lucrative position in psychology develops. The college teacher of psychology should be a person who has been trained to teach psychology. The calibre of his teaching should be high. It seems that in psychology this is even more urgent than in certain other subjects about which there are fewer misconceptions and which the student admits ignorance of more readily. The subject matter of the science of human behavior has become so vast and so complex in recent years that the need for learning how to present the ever-growing body of knowledge effectively is greater than ever. It seems obvious that graduate departments ought to offer some preparation for teaching, but most graduate schools to date have almost completely overlooked their responsibility in this regard. Yet these same schools want well-trained students from the very colleges where they send their inexperienced graduates to teach. It might be well for graduate departments to check the roster of their alumni to determine how many of these alumni are now in teach-

---

*Received in the Editorial Office on February 27, 1956.

ing posts in the nation's colleges and universities and to consider in what specific ways they feel they have prepared them for their posts, besides giving them content courses. Such a check on the vocational aspects of the graduate program in psychology is as important for teacher preparation as for preparation of counselors, clinical workers, and research workers.

No graduate department of psychology would claim today that mere reception of the master's degree or doctorate in psychology makes a person a clinician or counselor without training in special skills. The graduate school rightly concerns itself with offering adequate preparation to clinicians, counselors and the like. It provides courses in clinical practice, counseling techniques, and research methods but in most instances it fails to offer any help or direct preparation to prospective teachers. The fact that an individual has received the M.A. or Ph.D. seems to be considered adequate preparation for the potential teacher according to the present standards of most graduate departments of psychology. Most of us know better than to subscribe to such a belief.

A considerable number of the students who attend graduate schools upon attainment of their degrees do some college teaching, either full time or part time. Why do most graduate schools not assume responsibility for the preparation of these future teachers of psychology except in a very remote or indirect way? Why do those who want to teach not get some special training, even if not as intensive, as that of their colleagues who want to be counselors and clinicians? There seems to be a tacit assumption on the part of administrators of graduate departments of psychology that their students who have sat in the lecture halls of learned professors will automatically acquire the best teaching techniques and thus become good teachers of psychology themselves. While it is true that sometimes the student may absorb and imitate the good teaching qualities of his graduate professors, he still finds himself stranded and in a predicament when confronted with the concrete problems of planning courses, selecting textbooks, constructing examinations, and arranging demonstrations and field trips. Occasionally such a new teacher can call upon a trusted friend at his old graduate school. Usually he hesitates to admit his inadequacy at the institution where he has accepted a job.

Very frequently the newcomer to teaching carries to his college class the techniques of graduate-level instruction which have been his immediate experience. He fails to amend these for his less mature students. He crams far too much into single lectures. He dwells too much on his own specialty and may fail to present adequately the over-all essentials of his course. He is often totally inept at organizing his material and resorts to using the notes from his recently taken graduate course, offering his students a mere digest of that material. Sometimes, too, the young college teacher may make his teaching incidental to his own research, using the classroom as a kind of recruiting station for subjects to participate in his experiments. It is obvious that in the face of these practices the students are being cheated. What may not always be so obvious is that psychology as a science in the college curriculum also suffers. The student exposed to poor teaching receives an unfavorable impression not only of the teacher but often inevitably of the subject which was presented by that teacher. Anyone who has taught psychology at the college level knows how much more critical students are of bad teaching practices in their psychology teachers than in other subject matter teachers. Probably all of us have heard the old student refrain, "He's a psychologist. He ought to know better." The college classroom is an important locus for educating our citizenry not only in the contents of psychology, but its usefulness and value. The college teacher of psychology as a teacher, psychologist, and as a person should take this rôle as public relations officer for his science seriously and should use his position wisely to inculcate an appreciation for psychology as a science and a profession in his students.

Of course it must be admitted that teachers of psychology as well as teachers of other subjects are entitled to their idiosyncrasies, and surely it must be granted that, personality-wise, some people will be more effective teachers and some teachers will win more friends for their subject and influence more people than others. But in spite of this acknowledged variability as to personality, better preparation of psychology teachers would be an improvement. Such training can easily be provided. Certain basic information, skills, techniques, and principles of teaching psychology can and should be made part of the repertoire of the individual who wants to teach. The college which hires him should not have to begin to give him basic training. The graduate school could and should do its part by offering at least a seminar or short course on problems of teaching psychology. Such courses would help to reduce the numbers of inadequately prepared teachers of psychology launched by the graduate schools into the nation's colleges annually.

In the last few years colleges have been undertaking the job of training their new teachers of practically all subject matter areas. With the assistance of the Ford Foundation valuable seminars are being held in some institutions to train the newcomers to the college teaching profession. Preliminary training in techniques and methods could make these seminars and this in-service training even more valuable in the specific institutions where they are being conducted. The graduate departments of psychology should come to see and admit their deficiency in teacher training and take steps tc remedy this in the interests of psychology as a science and as a profession. In this way, too, the teaching of psychology will be dignified as a profession along with the other more practical pursuits in psychology. It might also serve to dispel the rather fallacious notion that anybody can teach. With increased dignity and recognition accorded to the teaching of psychology it might become possible to attract more competent and more zealous teachers who are needed now and for whom there will be an even greater need if college populations soar to the anticipated levels in the next few decades. More effective teaching should result in increased understanding of and appreciation for psychological science.

*Department of Psychology*
*Hunter College*
*New York, N. Y.*

*The Journal of Genetic Psychology,* 1957, **90**, 275-284.

# BOOKS

Now that there is a special APA journal completely devoted to the publication of book reviews, it is no longer necessary that other journals emphasize such publication. It has always been our conviction that book reviews are a low order of publication unless they carry information that is even more important than the book. However, the publication of book titles is a very important service, and we shall continue to render that service.

In any given issue of this journal, we may continue to publish one or more book reviews, but we do not consider such publication a major function of this journal. In line with this policy, we can no longer pay for such manuscripts.

## BOOKS RECENTLY PUBLISHED

BERNARD, H. W. Adolescent Development in American Culture. Yonkers-on-Hudson: World Book, 1957. Pp. 644.

WELSH, C. S., & DAHLSTROM, W. G., *Eds.* Basic Readings on the *MMPI* in Psychology and Medicine. Minneapolis: Univ. Minnesota Press, 1956. Pp. 656.

GRAHAM, E. C., & MULLEN, M. M. Rehabilitation Literature 1950-1955. New York: McGraw-Hill, 1956. Pp. 621.

SCHAFFNER, B., *Ed.* Group Processes. New York: Macy, 1956. Pp. 255.

MURPHY, G., *Ed.* A Unesco Study of Social Tensions in Aligarh. Ahmedabad: New Order Book, 1955. Pp. 206.

BARNOUW, E. Mass Communication. New York: Rinehart, 1956. Pp. 280.

GITTLER, J. B., *Ed.* Understanding Minority Groups. New York: Wiley, 1956. Pp. 139.

FEIGL, H., & SCRIVEN, M., *Eds.* Minnesota Studies in the Philosophy of Science. Minneapolis: Univ. Minnesota Press, 1956. Pp. 346.

JUNG, C. G. Symbols of Transformation (Collected Works 5). New York: Pantheon, 1956. Pp. 567.

FISHER, R. A. Statistical Methods and Scientific Inference. New York: Hafner, 1956. Pp. 173.

LEARY, T. Interpersonal Diagnosis of Personality. New York: Ronald Press, 1957. Pp. 518.

LIGON, E. H. Dimensions of Character. New York: Macmillan, 1956. Pp. 497.

NEWCOMB, W. W., JR. The Culture and Acculturation of the Delaware Indians. Ann Arbor: Univ. Michigan, 1956. Pp. 141.

DAVID, H. P., & VON BRACKEN, H. Perspectives in Personality Theory. New York: Basic Books, 1957. Pp. 435.

SCHREIBER, F. R. Your Child's Speech. New York: Putnam, 1956. Pp. 256.

MULLER, T. G. The Foundations of Human Behavior. New York: Putnam, 1956. Pp. 254.

NEUMANN, E. Amor and Psyche. New York: Pantheon, 1956. Pp. 181.

DOBZHANSKY, T. The Biological Basis of Human Freedom. New York: Columbia Univ. Press, 1956. Pp. 135.

VIRTUE, M. B. Family Cases in Court. Durham: Duke Univ. Press, 1956. Pp. 290.

ASHEY, W. R. An Introduction to Cybernetics. New York: Wiley, 1956. Pp. 295.

MAIER, N. R. F., SOLEM, A. R., & MAIER, A. A. Supervisory and Executive Development. New York: Wiley, 1957. Pp. 330.

FARNSWORTH, P. R., *Ed.* Annual Review of Psychology. Palo Alto: Annual Reviews, 1957. Pp. 502.

Buxton, C. E. College Teaching. New York: Harcourt, Brace, 1957. Pp. 404.
Levine, E. S. Youth in a Soundless World. New York: N. Y. Univ. Press, 1956. Pp. 217.
Cherry, C. On Human Communication. New York: Wiley, 1957. Pp. 333.
Havighurst, R. J., & Neugarten, B. L. Society and Education. Boston: Allyn & Bacon, 1957. Pp. 465.
Jersild, A. T. The Psychology of Adolescence. New York: Macmillan, 1957. Pp. 438.
Williams, R. J. Biochemical Individuality. New York: Wiley, 1956. Pp. 214.
Kutner, B., Fanshel, D., Togo, A. M., & Langner, T. S. Five Hundred Over Sixty. New York: Russell Sage, 1956. Pp. 345.
Ellis, A., & Brancale, R. The Psychology of Sex Offenders. Springfield: Thomas, 1956. Pp. 132.
Barbash, J. The Practice of Unionism. New York: Harper, 1956. Pp. 465.
Bursk, E. C., *Ed.* Human Relations for Management. New York: Harper, 1956. Pp. 372.
Martin, C. P. Psychology, Evolution and Sex. Springfield: Thomas, 1956. Pp. 166.
Krout, M. H., *Ed.* Psychology, Psychiatry and the Public Interest. Minneapolis: Univ. Minnesota Press, 1956. Pp. 217.
Wylie, L. Village in the Vaucluse. Cambridge: Harvard Univ. Press, 1957. Pp. 345.
Hall, C. S., & Lindzey, G. Theories of Personality. New York: Wiley, 1957. Pp. 572.
Simon, H. A. Models of Man. New York: Wiley, 1957. Pp. 287.
Hansen, H. After-Effects in the Behaviour of Mice. Copenhagen: Munksgaard, 1956. Pp. 423.
Foster, J. Sex Variant Women in Literature. New York: Vantage, 1956. Pp. 412.
Berrien, F. K., & Bash, W. H. Human Relations: Comments and Cases. New York: Harper, 1957. Pp. 564.
Underwood, B. J. Psychological Research. New York: Appleton-Century-Crofts, 1957. Pp. 298.
Frazier, E. F. The Negro in the United States. New York: Macmillan, 1957. Pp. 769.
Burtt, H. E. Applied Psychology. Englewood Cliffs: Prentice-Hall, 1957. Pp. 592.
Miner, J. B. Intelligence in the United States. New York: Springer, 1957. Pp. 180.
Carmichael, L. Basic Psychology. New York: Random House, 1957. Pp. 340.
Kornhauser, A., *Ed.* Problems of Power in American Democracy. Detroit: Wayne State Univ. Press, 1957. Pp. 239.
Johnson, R. I., Schalekamp, M., & Garrison, L. A. Communication. New York: McGraw-Hill, 1956. Pp. 361.
Rohde, A. R. The Sentence Completion Method. New York: Ronald Press, 1957. Pp. 301.
Langer, M. Learning to Live as a Widow. New York: Gilbert, 1957. Pp. 255.
Siegel, S. Nonparametric Statistics. New York: McGraw-Hill, 1956. Pp. 312.
Bois, J. S. Explorations in Awareness. New York: Harper, 1957. Pp. 212.
Bjerstedt, A. Interpretations of Sociometric Choice Status. Lund: Gleerup, 1956. Pp. 408.
Borgatta, E. F., & Meyer, H. J., *Eds.* Sociological Theory. New York: Knopf, 1956. Pp. 546.
Brameld, T. Toward a Reconstructed Philosophy of Education. New York: Dryden, 1956. Pp. 417.
Broudy, H. S., & Freel, E. L. Psychology for General Education. New York: Longmans, Green, 1956. Pp. 456.
Bugelski, B. R. The Psychology of Learning. New York: Holt, 1956. Pp. 523.
College Entrance Examination Board. Advanced Placement Program. New York: College Entrance Examination Board, 1956. Pp. 136.
Crow, L. D., & Crow, A. Understanding Our Behavior. New York: Knopf, 1956. Pp. 347.
Gardner, G. E., *Ed.* Case Studies in Childhood Emotional Disabilities. Vol. II. New York: American Orthopsychiatric Association, 1956. Pp. 368.

GOODENOUGH, F. L. Exceptional Children. New York: Appleton-Century-Crofts, 1956. Pp. 428.

GRAY, W. S., & ROGERS, B. Maturity in Reading: Its Nature and Appraisal. Chicago: Univ. Chicago Press, 1956. Pp. 273.

GRINKER, R. R. Toward a Unified Theory of Human Behavior. New York: Basic Books, 1956. Pp. 375.

HICKMAN, C. A., & KUHN, M. H. Individuals, Groups, and Economic Behavior. New York: Dryden, 1956. Pp. 266.

JORDAN, A. M. Educational Psychology: Growth and Learning. New York: Holt, 1956. Pp. 600.

LINDESMITH, A. R., & STRAUSS, A. L. Social Psychology. New York: Dryden, 1956. Pp. 703.

MARZOLF, S. S. Psychological Diagnosis and Counseling in the Schools. New York: Holt, 1956. Pp. 401.

MCCARTHY, J. L. Understanding Human Behavior. New York: Vantage, 1956. Pp. 258.

MCCLOSKEY, J. F., & COPPINGER, J. M., *Eds.* Operations Research for Management. Vol. II: Case Histories, Methods, Information Handling. Baltimore: Johns Hopkins, 1956. Pp. 563.

MCDANIEL, H. B., with SHAFTEL, G. A. Guidance in the Modern School. New York: Dryden, 1956. Pp. 526.

MEAD, M. New Lives for Old: Cultural Transformation—Manus, 1928-1953. New York: Morrow, 1956. Pp. 548.

MORENO, J. L., *Ed.* Sociometry and the Science of Man. Beacon, N. Y.: Beacon House, 1956. Pp. 474.

NIJMEGEN UNIVERSITY. PSYCHOLOGICAL CIRCLE. Psychology in Europe. Nijmegen and Utrecht: Dekker & Van de Vegt, 1956. Pp. 79.

ROSE, A. M. Sociology: The Study of Human Relations. New York: Knopf, 1956. Pp. 589.

RUESCH, J., & KEES, W. Nonverbal Communication: Notes on the Visual Perception of Human Relations. Berkeley: Univ. California Press, 1956. Pp. 205.

SARTRE, J. P. Being and Nothingness. New York: Philosophical, 1956. Pp. 638.

SMITH, C. Sensism: The Philosophy of the West. New York: Truth Seeker, 1956. Pp. 732 & 1612.

STEARN, J. Sisters of the Night: The Startling Story of Prostitution in New York Today. New York: Messner, 1956. Pp. 182.

STEPHENS, J. M. Educational Psychology: The Study of Educational Growth. New York: Holt, 1956. Pp. 717.

STEVENSON, G. S. Mental Health Planning for Social Action. New York: McGraw-Hill, 1956. Pp. 358.

STEWART, R. S., & WORKMAN, A. D. Children and Other People: Achieving Maturity Through Learning. New York: Dryden, 1956. Pp. 276.

WENGER, M. A., JONES, F. N., & JONES, M. H. Physiological Psychology. New York: Holt, 1956. Pp. 472.

WINSLOW, W. The Menninger Story. Garden City: Doubleday, 1956. Pp. 350.

ALBIG, W. Modern Public Opinion. New York: McGraw-Hill, 1956. Pp. 518.

ALLPORT, F. H. Theories of Perception and the Concept of Structure. New York: Wiley, 1955. Pp. 709.

BAUER, R. A., INKELES, A., & KLUCKHOHN, C. How the Soviet System Works. Cambridge: Harvard Univ. Press, 1956. Pp. 274.

BRODY, S. Patterns of Mothering: Maternal Influence During Infancy. New York: International Universities Press, 1956. Pp. 446.

COTTINGHAM, H. F. Guidance in Elementary Schools: Principles and Practices. Bloomington: McKnight, 1956. Pp. 325.

FRANK, L. K., & FRANK, M. Your Adolescent at Home and in School. New York: Viking, 1956. Pp. 336.

FROMM-REICHMANN, F., & MORENO, J. L., *Eds.* Progress in Psychotherapy. New York: Grune & Stratton, 1956. Pp. 352.

Gardner, E. F., & Thompson, G. G. Social Relations and Morale in Small Groups. New York: Appleton-Century-Crofts, 1956. Pp. 312.

Glover, E. On the Early Development of Mind. New York: Intern. Univ. Press, 1956. Pp. 483.

Halpin, A. W. The Leadership Behavior of School Superintendents. Columbus: College of Education, Ohio State Univ., 1956. Pp. 109.

Gottlieb, J. S., Cleghorn, R. A., Himwich, H. E., Kline, N. S., Llhamon, W. T., & Saslow, G. An Evaluation of the Newer Psychopharmacologic Agents and Their Rôle in Current Psychiatric Practice. Washington, D. C.: American Psychiatric Association, 1956. Pp. 129.

Harris, M. Town and Country in Brazil. New York: Columbia Univ. Press, 1956. Pp. 302.

Henry, W. E. The Analysis of Fantasy: The Thematic Apperception Technique in the Study of Personality. New York: Wiley, 1956. Pp. 305.

Hurlock, E. B. Child Development. New York: McGraw-Hill, 1956. Pp. 703.

Israel, J. Self-Evaluation and Rejection in Groups. Stockholm: Almqvist & Wiksell, 1956. Pp. 250.

Klopfer, B., *Ed.* Developments in the Rorschach Technique. Vol. II: Fields of Application. Yonkers-on-Hudson, N. Y.: World Book, 1956. Pp. 828.

Kornhauser, A., Mayer, A. J., & Sheppard, H. L. When Labor Votes: A Study of Auto Workers. New York: University Books, 1956. Pp. 352.

Kranefeldt, W. M. Therapeutische Psychologie: IHR Weg Durch Die Psychoanalyse. Berlin: Gruyter, 1956. Pp. 152.

Linschoten, J. Strukturanalyse der Binokularen Tiefenwahrnehmung. Groningen, Djarkarta: Wolters, 1956. Pp. 573.

Lippman, H. S. Treatment of the Child in Emotional Conflict. New York: Blakiston Div., McGraw-Hill, 1956. Pp. 298.

McCary, J. L., *Ed.* Psychology of Personality: Six Modern Approaches. New York: Logos, 1956. Pp. 383.

Michal-Smith, H. The Mentally Retarded Patient. Philadelphia: Lippincott, 1956. Pp. 203.

Moustakas, C. E., *Ed.* The Self: Explorations in Personal Growth. New York: Harper, 1956. Pp. 284.

Murphy, L. B., & Associates. Personality in Young Children. Vol. I & II. New York: Basic Books, 1956. Pp. 424 & 267.

Pavlov, I. P. Selected Works. Moscow: Foreign Languages Publishing House, 1955. Pp. 654.

Rado, S. Psychoanalysis of Behavior. New York: Grune & Stratton, 1956. Pp. 387.

Rado, S., & Daniels, G. E., *Eds.* Changing Concepts of Psychoanalytic Medicine. New York: Grune & Stratton, 1956. Pp. 248.

Blumenfeld, W. Introduccion a la Psicologia Esperimentale. Lima: Editorial Cultura Antartica S.A., 1954. Pp. 485.

Bossard, J. H. S., & Boll, E. S. The Large Family System: An Original Study in the Sociology of Family Behavior. Philadelphia: Univ. Pennsylvania Press, 1956. Pp. 325.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. A Study of Thinking. New York: Wiley, 1956. Pp. 330.

Duker, S., & Nally, T. The Truth About Your Child's Reading. New York: Crown, 1956. Pp. 181.

Fleishman, E. A., Harris, E. F., & Burtt, H. E. Leadership and Supervision in Industry. Columbus: Bureau of Educational Research, Ohio State Univ., 1955. Pp. 110.

Franklin, E. Tonality as a Basis for the Study of Musical Talent. Goteborg: Gumperts Forlag, 1956. Pp. 193.

Glueck, S., & Glueck, E. Physique and Delinquency. New York: Harper, 1956. Pp. 339.

Kracauer, S., & Berkman, P. L. Satellite Mentality: Political Attitudes and Propaganda Susceptibilities of Non-Communists in Hungary, Poland, and Czechoslovakia. New York: Praeger, 1956. Pp. 194.

Lesser, M. S., & Keane, V. R. Nurse-Patient Relationships in a Hospital Maternity Service. St. Louis: C. V. Mosby, 1956. Pp. 301.

Lewis, N. D. C., Landis, C., & King, H. E., *Eds.* Studies in Topectomy. New York: Grune & Stratton, 1956. Pp. 248.

Matte Blanco, I. Estudios de Psicologia Dinamica. Santiago: Univ. de Chile, 1955. Pp. 329.

Matte Blanco, I. Lo Psiquico y la Naturalezo Humana: Hacia un Planteamiento Experimental. Santiago: Univ. de Chile, 1954. Pp. 263.

McCord, W., & McCord, J. Psychopathy and Delinquency. New York: Grune & Stratton, 1956. Pp. 230.

Morris, C. Varieties of Human Value. Chicago: Univ. Chicago Press, 1956. Pp. 209.

Odier, C. Anxiety and Magic Thinking. New York: International Univ. Press, 1956. Pp. 302.

Pigors, P., & Myers, C. A. Personnel Administration: A Point of View and a Method. New York: McGraw-Hill, 1956. Pp. 711.

Radin, P. The Trickster: A Study in American Indian Mythology. New York: Philosophical, 1956. Pp. 211.

Roe, A. The Psychology of Occupations. New York: Wiley, 1956. Pp. 340.

Sechehaye, M. A New Psychotherapy in Schizophrenia: Relief of Frustrations by Symbolic Realization. New York: Grune & Stratton, 1956. Pp. 199.

Slotkin, J. S. The Peyote Religion: A Study in Indian-White Relations. Glencoe: Free Press, 1956. Pp. 195.

Smith, E. S. The Dynamics of Aging. New York: Norton, 1956. Pp. 191.

Spence, K. W. Behavior Theory and Conditioning. New Haven: Yale Univ. Press, 1956. Pp. 262.

Spiro, M. E. Kibbutz: Venture in Utopia. Cambridge: Harvard Univ. Press, 1956. Pp. 266.

Steiner, F. Taboo. New York: Philosophical, 1956. Pp. 154.

Steinfeld, J. I. A New Approach to Schizophrenia. New York: Merlin, 1956. Pp. 195.

Sullivan, H. S. Clinical Studies in Psychiatry. New York: Norton, 1956. Pp. 386.

Summers, M. The History of Witchcraft. New York: Univ. Books, 1956. Pp. 353.

Adams, G. S., & Torgerson, T. L. Measurement and Evaluation for the Secondary-School Teacher. New York: Dryden, 1956. Pp. 658.

Boulding, K. E. The Image. Ann Arbor: Univ. Michigan Press, 1956. Pp. 175.

Bursk, E. C., *Ed.* Human Relations for Management. New York: Harper, 1956. Pp. 372.

Eisenstein, V. W., *Ed.* Neurotic Interaction in Marriage. New York: Basic Books, 1956. Pp. 352.

Getz, S. Environment and the Deaf Child. Springfield: Thomas, 1956. Pp. 173.

Hirning, J. L., & Hirning, A. L. Marriage Adjustment. New York: American Book, 1956. Pp. 456.

Hofstaetter, P. R. Sozialpsychologie. Berlin: Gruyter, 1956. Pp. 181.

Hunt, W. A. The Clinical Psychologist. Springfield: Thomas, 1956. Pp. 206.

Kaplan, B., & Plaut, T. F. A. Personality in a Communal Society: An Analysis of the Mental Health of the Hutterites. Lawrence: Univ. Kansas, 1956. Pp. 116.

Krout, M. H., *Ed.* Psychology, Psychiatry and the Public Interest. Minneapolis: Univ. Minnesota Press, 1956. Pp. 217.

Liebman, S., *Ed.* Management of Emotional Problems in Medical Practice. Philadelphia: Lippincott, 1956. Pp. 152.

Mannoni, O. Prospero and Caliban: The Psychology of Colonization. New York: Praeger, 1956. Pp. 218.

Mezer, R. R. Dynamic Psychiatry in Simple Terms. New York: Springer, 1956. Pp. 174.

Nyswander, M. The Drug Addict as a Patient. New York: Grune & Stratton, 1956. Pp. 179.

*(Henry, W. E. The Analysis of Fantasy: The Thematic Apperception Technique in the Study of Personality. New York: Wiley, 1956. Pp. 305.)*

REVIEWED BY ROBERT M. ALLEN

The publisher has added to a distinguished series of volumes in the field of the assessment of human behavior. This book is not intended primarily as a manual of administration even though suggestions are contained in a brief appendix. However, the discussion of the administration technique yields to the basic theme of the author's client-centered approach. Note is also made of the examiner's rôle in the evaluative process.

By now it must be obvious to psychologists that the influence of the examiner has a proper place in assessing the usefulness of a projective device. This issue has been raised and discussed at length by Lord (4), Sander and Cleveland (7), and Allen (1) in connection with the Rorschach Test. Additional theoretical and pertinent comments have been offered in great detail by Sarason (8) and Schafer (9). There can be little doubt as to the extent to which examiner hostility, needs, and press play a rôle in the dynamics of the testing situation to influence responsiveness, scoring, and interpretation. These may be subtle and unintentional variables, but they do exist in many instances. The situation becomes all the more contaminated when the examiner quite deliberately imposes himself onto the situation beyond his rôle of producing the proper stimuli, observing, and recording.

With particular reference to the thematic test, Henry has two excellent suggestions for minimizing examiner-influence. First, by exemplifying in his own interpretations the maxim of keeping ". . . interpretation close to data," i.e., ". . . by the close association of evidence and inference:" and second, by stating his preference for formulating theoretical bases for understanding the normal personality. The latter cannot help but add to the comprehension of the non-normal. The research aspect may not have immediate interest for the busy clinician, but its contribution to personality theory and to test theory cannot be overemphasized.

The author includes an excellent bibliography of 568 references divided into pertinent topics ranging from: The Test and Its Manuals: Administration, Scoring, and Interpretation, to Closely Related Techniques, Including Adaptation for Special Groups. Other areas covered in the bibliography are theoretical, methodological, and normative studies. For the beginning and advanced student the categorical subject-related bibliography could most certainly serve as the beginning focus for a reading and literature research program in the thematic technique and in the projective procedure.

The formal structure of this book is divided into three parts: the interpretive process, analysis of single cases, and a discussion of the Thematic Apperception Test (*TAT*) pictures. The basic rationale for the thematic approach to personality assessment is rooted, according to Henry, in the ". . . interaction of individual personality and the stimulation represented by the pictures" (3, p. vii). This, in turn, leads to inferences regarding the dynamics of the observed behavior—a higher order of abstraction that is anchored in the individual's perception of his internal and external world, directly and through its symbols.

Henry sets the limits to interpretive flights which exposes to view his own cautious bias and his awareness of what is taking place: "First," he writes, "I have at all times attempted to keep interpretation close to data, feeling that progress in research in this field will benefit most by the close association of evidence and inference. To this end, I have deliberately restrained myself in the face of the temptation to make secondary interpretations based upon first-order interpretations" (3, p. viii). He does not assume that even these first-order inferences are equally valid for the subjects. The central problem of validity is not answered in this volume—especially since it is difficult, at times, to completely verify the closeness or distanciation of the testee's fantasies to their dynamic well-springs in real life. This is a refreshing attitude.

While the title of this volume limits, presumably, the contents to the *TAT*, the interpretive principles can easily be applied to the other projective techniques. For example, Henry writes (3, p. 9): The Thematic Apperception Technique is a device for the analysis of (these) underlying psychological needs and feelings and the ways in which they are handled by the individual." Certainly any attempts to handle the Rorschach protocol or the figure-drawing product in terms other than those emphasized in this statement would result either in untenable, albeit interesting, flights of fancy or in sterile formalized deduction from still more sterile formulae. Neither are desirable nor helpful to the testee.

To Henry all that passes for fantasy is related to the life space of the individual—this includes not only the distanciated and the distorted but also all that seems quite normal and acceptable. This point of view should be taken seriously if clinical psychologists and other users of projective devices are to minimize the general charge that all persons are normal until they take a projective test.

Henry introduces a basic assumption into the interpretive process—constant emotional movement which reduces to a leavening or intervening variable which helps ". . . to explain the interrelationships of behavior and motive and the continuity and consistency of personality through time" (3, p. 19). The essential terms attribute behavior to motivation and the need to maintain emotional homeostasis as the person copes with internal and external stimuli. Like other similarly oriented theories, the shortcomings become noticeable in the efforts to deal with the issue of interactionism, the leaven that holds together behavior, motive, set, adaptation, and the person.

The elaboration of the formal and content aspects of the *TAT* plates as cues to interpretation is a reminder of Piotrowski's (6) general principles of thematic interpretation with all of their cautions for the wary clinician. Throughout the volume there runs the theme of relating the story to inner dynamics as mirrors of the testee's perception of his world and attempts at adjustment to it. Guides are set up for the analysis of the protocol in eight areas: mental approach, imaginative, processes, family dynamics, inner adjustment, emotional reactivity, sexual adjustment, behavioral approach, and descriptive and integrative summary. It is in the last area that the transcendental quality of the interpretation becomes a highly individual matter with each test interpreter.

Each area is elaborated in detail with illustrations from selected portions of stories related by a variety of subjects to the different pictures of the *TAT* series. The reader cannot help but be impressed with the client-centeredness of the entire discussion of interpretive principles. Henry makes every effort to emphasize concern for understanding the testee as part of a behavior-determined phenomenon in a social setting. This principle is closely adhered to in the illustrative analyses in the second portion of the book.

Psychologists who are familiar with Henry's earlier monograph (2) will recognize his point of view as being continued and elaborated upon in the present book. The Content Characteristics of the subject's stories are the foci of thematic analysis and interpretation. Along with others (10, 13), Henry appreciates the futility of quantification as a method of teasing out the components of the subject's thema. This is especially so when the stories are compounded in terms of their subtler nuances with consequent increase in story-categories and subcategories. The global approach is still the major emphasis with qualitative analysis of the thematic form and content as the basic materials. The variables that comprise these two analytic categories are discussed in detail in Chapters 4 and 5 of this volume. Numbers as such play no part in the process of thematic analysis and interpretation.

Actually, Henry's concepts of fantasy analysis is one of several such schema. It is a departure from Murray's (5) need-press analysis and its variants in Stein's (10) and Tompkins' (11) systems. It comes closest to Wyatt's (12) mode of story analysis in terms of 15 variables. Wyatt does suggest, however, that in his scheme ". . . a certain amount of numerical scoring is possible . . ." (12, p. 329). It may be assumed that this is recognition of the fact that individual and group trends, the latter as norms, may be established. How this could contribute to the objectivity of the thematic procedure does not seem to be a topic of discussion.

Part 2 of this volume is given over to a presentation of four protocols and their analysis. The first two are of Joe, a normal 14-year-old youngster, and of Carl, a 29-year-old immigrant. The second two reflect Henry's particular research interest: Charlie, a 14-year-old Dakota Sioux subject and an 11-year-old Navaho girl, Betsy. In each instance the verbatim *TAT* protocols are followed by detailed sequential analyses for each picture plus excerpts from each subject's autobiography or life history. Where available, Rorschach Test data are also given.

The presentation in Part 3 of each *TAT* picture with its predetermined manifest and latent stimulus demands upon the subject and the most frequently given plots flirts with the danger of setting up interpretive criteria. The weakness inherent in this approach is the very one which Henry is attempting to caution against—equating responses with check-list type of interpretation. However, Henry's point of view is that each picture does have certain "given" features which make demands upon the testee, i.e., to elicit responses because of their attention-pull qualities. These contribute to the popularity of certain response-contents and the resultant theme. In turn, these help in setting up referents for the discovery of distortions, misperceptions, and the range of "normal variations" as clues to aberrant and adjusted personality makeup. Despite the inherent danger, interpretation must have a starting point. This weakness is ascribable to projective devices other than the *TAT* (1). Reliability and validity data are not addressed in this volume. Very little attention is given to the statistical aspect of the *TAT*.

Henry has succeeded in creating a reader-atmosphere of imparting comfort in the use of a probing projective assessment procedure. The pages of this book turn easily. In sum, this volume is a constructive contribution to the clinical psychologist's repertory of methods for using a significant tool effectively.

## REFERENCES

1. ALLEN, R. M. Elements of Rorschach Interpretation. New York: International Univ. Press, 1954.
2. HENRY, W. E. The thematic apperception technique in the study of culture-personality relations. *Genet. Psychol. Monog.*, 1947, **35**, 3-135.
3. ———. The Analysis of Fantasy. New York: Wiley, 1956.
4. LORD, E. Experimentally induced variations in Rorschach performance. *Psychol. Monog.*, 1950, **64**. Pp. 31.
5. MURRAY, H. A. Thematic Apperception Test Manual. Cambridge: Harvard Univ. Press, 1943.
6. PIOTROWSKI, Z. A. A new evaluation of the Thematic Apperception Test. *Psychoanal. Rev.*, 1950, **37**, 101-127.
7. SANDERS, R., & CLEVELAND, S. E. The relationship between certain examiner personality variables and subjects' Rorschach scores. *J. Proj. Tech.*, 1953, **17**, 34-50.
8. SARASON, S. B. The Clinical Interaction. New York: Harper, 1954.
9. SCHAFER, R. Psychoanalytic Interpretation in Rorschach Testing. New York: Grune & Stratton, 1954.
10. STEIN, M. I. The *TAT*. Cambridge: Addison-Wesley Press, 1948.
11. TOMPKINS, S. The Thematic Apperception Test. New York: Grune & Stratton, 1947.
12. WYATT, F. The scoring and analysis of the Thematic Apperception Test. *J. of Psychol.*, 1947, **24**, 319-330.

*Department of Psychology*
*University of Miami*
*Coral Gables 46, Florida*

BUREAU OF EDU[illegible] ARCH.